

ON THE FORMATION OF PLANETESIMALS VIA SECULAR GRAVITATIONAL INSTABILITIES WITH
TURBULENT STIRRING

ANDREW N. YODIN

Harvard Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

ApJ, accepted

ABSTRACT

We study the gravitational instability (GI) of small solids in a gas disk as a mechanism to form planetesimals. Dissipation from gas drag introduces secular GI, which proceeds even when standard GI criteria for a critical density or Toomre's Q predict stability. We include the stabilizing effects of turbulent diffusion, which suppresses small scale GI. The radially wide rings that do collapse contain up to ~ 0.1 Earth masses of solids. Subsequent fragmentation of the ring (not modeled here) would produce a clan of chemically homogenous planetesimals. Particle radial drift time scales (and, to a lesser extent, disk lifetimes and sizes) restrict the viability of secular GI to disks with weak turbulent diffusion, characterized by $\alpha \lesssim 10^{-4}$. Thus midplane dead zones are a preferred environment. Large solids with radii $\gtrsim 10$ cm collapse most rapidly because they partially decouple from the gas disk. Smaller solids, even below \sim mm-sizes could collapse if particle-driven turbulence is weakened by either localized pressure maxima or super-Solar metallicity. Comparison with simulations that include particle clumping by the streaming instability shows that our linear model underpredicts rapid, small scale gravitational collapse. Thus the inclusion of more detailed gas dynamics promotes the formation of planetesimals. We discuss relevant constraints from Solar System and accretion disk observations.

Subject headings: hydrodynamics — instabilities — planetary systems: formation — planetary systems: protoplanetary disks — solar system: formation — turbulence

1. INTRODUCTION

The discovery of a diverse population of extrasolar planets¹ motivates the development of planet formation theories that are (at least) physically consistent and (at best) predictive of the rapidly growing data set. Forming planetesimals above kilometer sizes is both a crucial first step and testable in our own Solar system. Chiang & Youdin (2010, hereafter CY10) and Youdin (2010) review the main physical processes and theoretical puzzles.

A promising route to planetesimal formation is the clumping of solids by the streaming instability (SI) in protoplanetary disks (Youdin & Goodman 2005; Youdin & Johansen 2007; Johansen & Youdin 2007), which then induces gravitational collapse (Johansen et al. 2007, 2009b). SI is a drag instability (Goodman & Pindor 2000) that arises from the aerodynamic back reaction of solids on gas, and does not require self-gravity.

Strong SI-induced clumping appears to require fairly large solids with $\tau_s \sim 1$, where

$$\tau_s \equiv \Omega t_{\text{stop}} \quad (1)$$

measures the dynamical relevance of drag forces with t_{stop} the time scale for damping particle relative motion and Ω the Keplerian orbital frequency. Thus τ_s increases with particle radius a . In standard disk models, marginal coupling, $\tau_s = 1$, corresponds to $a \lesssim 1$ meter near 1 AU, and smaller sizes in the outer disk where gas densities are lower (see Fig. 1). For small solids with $\tau_s \ll 1$, SI operates but it stirs particles

without producing strong clumping (Johansen & Youdin 2007; Bai & Stone 2010a). Most other particle concentration mechanisms also become inefficient for $\tau_s \ll 1$ (see CY10, Youdin 2010).

It is desirable to find a planetesimal formation mechanism that operates when $\tau_s \ll 1$. Protoplanetary disks initially contain sub-micron-sized grains. Collisional sticking alone appears unlikely to produce meter-sized bodies (Benz 2000; Youdin & Shu 2002; Sekiya & Takeda 2003; Youdin 2004; Zsom et al. 2010). Furthermore, many primitive meteorites are composed primarily of chondrules, inclusions with $a \sim 0.1 - 1$ mm (Russell et al. 2006). Cuzzi et al. (2008) propose that turbulent eddies concentrate chondrules into clumps that become 10 — 100 km planetesimals. As discussed more extensively in CY10, this concentration mechanism normally applies to much smaller scales. More study of mechanisms to concentrate $\tau_s \ll 1$ solids is needed.

This work considers gravitational instabilities (GI) that are secular, meaning mediated by dissipation in the form of gas drag. The original proposals that planetesimals formed by GI (Safronov 1969; Goldreich & Ward 1973) applied the standard, dissipationless GI criterion of Toomre (1964). This approach gives the classic result of \sim kilometer-sized planetesimals. The fact that gas drag affects GI should not be surprising, especially for $\tau_s \ll 1$. Section 6.5 summarizes previous work on secular GI.

This paper contains an improved treatment of gas turbulence, notably the stabilizing effects of turbulent diffusion. We use the turbulent stirring model of Youdin & Lithwick (2007), which includes the effects of orbital oscillations. Our simplified treatment of gas dynamics does not capture SI or other non-gravitational concentration mechanisms. However, we do vary the disk

ayoudin@cfa.harvard.edu

¹ See <http://exoplanet.eu>, compiled by Jean Schneider, for an updated catalog with references.

“metallicity”

$$Z \equiv \Sigma/\Sigma_g, \quad (2)$$

the ratio of particle (unsubscripted) to gas surface densities. Large Z values mimic particle concentration or gas depletion on large scales (compared to the wavelength) and accelerate secular GI.

We emphasize one general result from our analysis. The critical or Roche-like² density

$$\rho_R \simeq 0.6 \frac{M_*}{R^3} \quad (3)$$

is *not* a useful discriminant for gravitational collapse when drag forces are significant, as they are for planetesimal formation. We use M_* for the central stellar mass and R for the radial distance from that star. Dissipation allows collapse to proceed for particle densities $\rho \ll \rho_R$. We prove this result in §3.1.3 and further demonstrate it in Fig. 6. Moreover, if $\rho > \rho_R$ in only a small volume, then collapse is avoidable. Gas drag limits collapse speeds to the terminal velocity, and turbulent diffusion smoothes small scale density perturbations.

Since secular GI can require many orbits, it is restricted by radial drift speeds and ultimately disk lifetimes. Observations find protoplanetary disks with ages up to several million years (Hernández et al. 2007). Planetesimal formation within 1 Myr should still allow time for the growth of gas and ice giants by core accretion (Lissauer & Stevenson 2007). The existence of primitive meteoritic inclusions with age spreads of several Myr (Russell et al. 2006) supports a longer time scale for planetesimal formation in the inner Solar System.

This article is organized as follows. We start with an order of magnitude derivation of secular GI in §2. We present our full model in §3 including: a derivation of the dispersion relation in §3.1, the turbulent stirring model in §3.2 and the protoplanetary disk model in §3.3. Section 4 gives the main numerical results covering: the variation of GI properties with disk parameters in §4.2 and a calculation of the strongest levels of turbulence that allow collapse in §4.3. Section 5 explores the analytic behavior of secular secular GI and the transition to standard GI. We discuss several aspects of our results in §6 including a comparison to simulations in §6.1, Solar System constraints in §6.2 and metallicity thresholds in §6.3. Section 7 concludes with a brief summary. Appendix A addresses turbulent drag laws and the behavior of secular GI in less relevant regions of parameter space is relegated to Appendix B.

2. SECULAR COLLAPSE OF SMALL SOLIDS

We begin with an order-of-magnitude derivation of the main result of this paper: the gravitational collapse of small solids in a gas disk. Small means that drag forces are strong with $\tau_s \ll 1$. Though probably less relevant for planetesimal formation, secular GI also exists for $\tau_s \gg 1$, see §5.2.

Consider a ring-like surface density perturbation of amplitude σ and radial length scale λ . The radial gravity

² The actual Roche density refers to tidal disruption of a sphere and is numerically larger (see CY10). The numerical coefficient in equation (3) is from the buckling modes of Sekiya (1983), which we discuss further in §6.5.

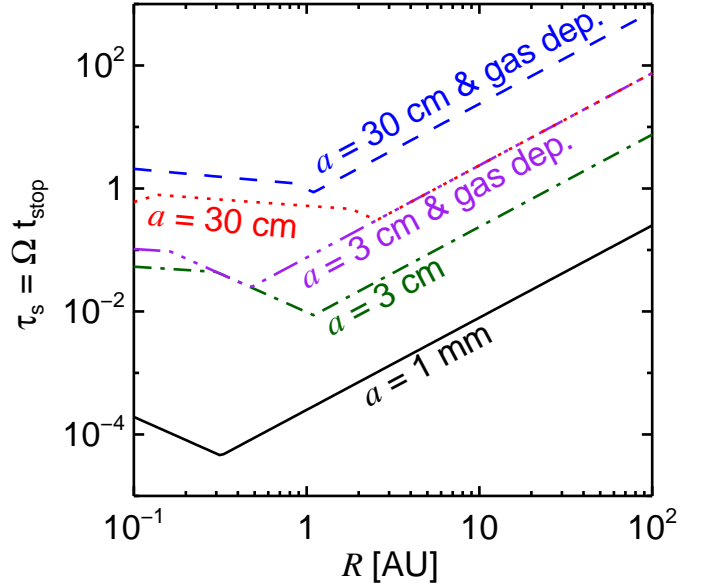


Figure 1. Aerodynamic stopping times (normalized by the orbital frequency) vs. radial distance in the disk midplane. For our reference minimum mass model ($F = 1$), we plot τ_s for particle sizes: $a = 1$ mm (black solid curve), $a = 3$ cm (green dot-dashed curve) and $a = 30$ cm (red dotted curve). For a gas depleted disk ($F = 0.1$) we again plot τ_s for $a = 3$ cm (purple tripe dot-dashed curve) and $a = 30$ cm (blue dashed curve). The kinks in the curves correspond to transitions between drag laws, as described in §3.3.2 and appendix A.

at the edge of the perturbation is $g_R \sim G\sigma$. If this result is unfamiliar, recall that the acceleration towards an infinitely long wire of linear mass density Λ is $2G\Lambda/\lambda$ at a distance λ . Then identify $\sigma \sim \Lambda/\lambda$ as if the ring were concentrated in a thin wire, which can be approximated as straight and infinitely long because $\lambda \ll R$, the radial distance to the star.

The particles’ response to self-gravity is mediated by gas drag. Radial sedimentation to the center of the ring occurs at the terminal velocity

$$u = g_R t_{\text{stop}} \sim G\sigma t_{\text{stop}}. \quad (4)$$

It is reasonable to ask why the radial gravity induces radial collapse instead of merely shifting the orbital speed by adjusting centripetal balance. The reason is that the drag on this extra orbital motion would give radial infall that exceeds the terminal velocity limit, hence equation (4) applies. This explanation assumes the pressure-supported gas remains stationary in response to the particle perturbations, a critical issue that we discuss further in §6.5.

By mass continuity, the amplitude of the density perturbation grows at a rate

$$s \sim \frac{\Sigma u}{\sigma \lambda} \sim \frac{G\Sigma t_{\text{stop}}}{\lambda}. \quad (5)$$

To understand whether and how fast collapse occurs, we must consider stabilizing influences that restrict the allowable λ . The excess angular momentum in a perturbation is removed very rapidly, in $t_{\text{stop}} \ll 1/s$ as we can verify shortly. Thus there is no obstacle to long wave-

length collapse, and instability is guaranteed. We now consider impediments to small scale collapse that set the minimum λ .

2.1. Effective Pressure Support: Less Relevant

We first consider the stabilization by the effective pressure of particle random velocities, c . Gravity, g_R exceeds the effective pressure acceleration, $c^2\sigma/(\Sigma\lambda)$, for wavelengths above

$$\lambda_P \sim \frac{c^2}{G\Sigma}. \quad (6)$$

From equation (5) the fastest growth rate allowed by effective pressure support is

$$s_P \sim \left(\frac{G\Sigma}{c}\right)^2 t_{\text{stop}} \sim \Omega \frac{\tau_s}{Q_T^2}. \quad (7)$$

$Q_T = c\Omega/(\pi G\Sigma)$ is the standard Toomre (1964) gravitational stability parameter, but unlike standard GI we have no requirement that $Q_T < 1$.

2.2. Diffusive Support: Dominant

Next we consider the stabilizing effect of particle diffusivity, D , due to stirring by gas turbulence. We express

$$D \approx c^2/\Omega. \quad (8)$$

as the product of the particle speed, c , and path length, c/Ω , which arises because disk turbulence randomizes the particle trajectory in an eddy turnover time $t_{\text{eddy}} = 1/\Omega$. See §3.2 for more details on turbulent stirring.

For the collapse rate of equation (5) to outpace the diffusive spreading time scale λ^2/D , wavelengths must exceed

$$\lambda_D \sim \frac{D}{G\Sigma t_{\text{stop}}} \sim \frac{\lambda_P}{\tau_s}. \quad (9)$$

Thus $\lambda_D \gg \lambda_P$, by our assumption of strong drag. Diffusion is the more effective stabilizing influence that sets the actual minimum λ .

Secular GI thus has a collapse rate

$$s \sim \frac{(G\Sigma t_{\text{stop}})^2}{D} \sim \Omega \left(\frac{\tau_s}{Q_T}\right)^2, \quad (10)$$

that is slower than s_P by $\tau_s \ll 1$. Growth becomes significantly slower as particle size, and thus τ_s , declines.

Previous studies of dissipative collapse that neglected turbulent diffusion (Ward 1976, 2000; Youdin 2005a,b) were thus overly optimistic. However the qualitative result that collapse does not require $Q_T < 1$ remains.

Armed with this basic description of the instability, some readers may wish to skip ahead to the numerical results of §4 and the figures. To better understand the model ingredient and caveats, we do recommend coming back to §3.

3. FORMULATION OF THE PROBLEM

3.1. Dynamical Equations

We consider the local, linear dynamics of solid particles in a protoplanetary disk subject to gas drag, turbulent mixing and self-gravity. We adopt the local shearing box model for a patch of the disk (Goldreich & Lynden-Bell

1965) with a frame centered on a fiducial radius that rotates with the local Kepler frequency Ω . The local radial and azimuthal coordinates are x and y , respectively. We decompose the particle surface density $\Sigma_p = \Sigma + \sigma$ and velocity fields $\mathbf{V} = \mathbf{V}_o + u\hat{x} + v\hat{y}$ into time-steady backgrounds — constant Σ and the Keplerian shear $\mathbf{V}_o = -(3/2)\Omega x\hat{y}$ — plus small amplitude perturbations. The perturbations are axisymmetric as appropriate for slow motions stretched by shear. Evolution in time, t , follows the linearized, height-integrated equations of continuity and force:

$$\frac{\partial\sigma}{\partial t} + \Sigma \frac{\partial u}{\partial x} = D \frac{\partial^2\sigma}{\partial x^2}, \quad (11a)$$

$$\frac{\partial u}{\partial t} - 2\Omega v = -\frac{u}{t_{\text{stop}}} - \frac{c^2}{\Sigma} \frac{\partial\sigma}{\partial x} - \frac{\partial\phi}{\partial x}, \quad (11b)$$

$$\frac{\partial v}{\partial t} + \Omega u/2 = -\frac{v}{t_{\text{stop}}}, \quad (11c)$$

The right hand side of equation (11a) describes mass diffusion, D , induced by turbulent gas. The effects of this term have also been considered by Shariff & Cuzzi (personal communication). The first term on the RHS of equations (11b) and (11c) are drag forces on particle motion relative to the static gas background. The aerodynamic stopping time, t_{stop} is evaluated in terms of particle sizes and gas densities in §3.3.2. The last two terms on the RHS of equation (11b) are standard in analyses of gravitational instability (GI): the effective pressure acceleration from (isothermal) random particle velocities, c , and the acceleration from the disk's self-gravitational potential ϕ . The validity of these model equations is discussed in §6.4.

We give the solution to Poisson's equation in terms of Fourier amplitudes which are denoted by tildes and have a spatiotemporal dependence $\propto \exp(st + ikx)$:

$$\tilde{\phi} = -\frac{2\pi G\tilde{\sigma}}{k} \mathcal{T}(kh), \quad (12)$$

The softening term, $\mathcal{T}(kh) = 1/(1 + kh)$ mimics the effects of finite particle layer thickness, h (Shu 1984). This factor connects the thin disk potential of long waves, $kh \ll 1$, to the three dimensional solution for $kh \gg 1$. While the modes studied in this paper are generally long, \mathcal{T} insures against the appearance of artificially short waves in a height-integrated model.

3.1.1. Dimensionless Dispersion Relation

We use the standard GI wavelength

$$\lambda_G \equiv 2\pi\ell_G \equiv 2\pi^2 G\Sigma/\Omega^2 \quad (13)$$

and the orbital time, $1/\Omega$, to define dimensionless velocities, $\{U, V\} \equiv \{u, v\}/(\ell_G\Omega)$, radial position $X \equiv x/\ell_G$ and time $T = \Omega t$. The growth rate, s , and wavenumber, $k = 2\pi/\lambda$, are non-dimensionalized as

$$\mathcal{K} \equiv k \cdot \ell_G = \lambda_G/\lambda, \quad (14a)$$

$$\gamma = s/\Omega, \quad (14b)$$

so the Fourier dependence becomes: $\exp(i\mathcal{K}X + \gamma T)$. In terms of dimensionless Fourier amplitudes, the equations of motion read

$$(\gamma + \mathcal{K}^2 Q_D^2) \tilde{\sigma}/\Sigma + i\mathcal{K}\tilde{U} = 0, \quad (15a)$$

$$(\gamma + 1/\tau_s)\tilde{U} - 2\tilde{V} = i\frac{1 - \mathcal{F}_W}{\mathcal{K}}\frac{\tilde{\sigma}}{\Sigma}, \quad (15b)$$

$$(\gamma + 1/\tau_s)\tilde{V} + \tilde{U}/2 = 0. \quad (15c)$$

The cubic dispersion relation for the (possibly complex) growth rate $\gamma(\mathcal{K})$ is

$$\left[(\gamma + \tau_s^{-1})^2 + 1\right](\gamma + \mathcal{K}^2 Q_D^2) = (\gamma + \tau_s^{-1})(1 - \mathcal{F}_W). \quad (16)$$

We follow Ward (1976, 2000) in parameterizing the strength of self-gravity with the function

$$\mathcal{F}_W \equiv \frac{\Omega^2 - 2\pi G \Sigma k \mathcal{T}(kh) + k^2 c^2}{\Omega^2} \quad (17a)$$

$$= 1 - \frac{2\mathcal{K}}{1 + \mathcal{K}Q_R} + Q_T^2 \mathcal{K}^2. \quad (17b)$$

Smaller \mathcal{F}_W indicates stronger self gravity.

At this point, the model requires four input parameters. These are the stopping time τ_s plus three gravitational stability parameters;

$$Q_T \equiv \frac{c\Omega}{\pi G \Sigma}, \quad (18a)$$

$$Q_R \equiv \frac{h\Omega^2}{\pi G \Sigma} \approx 0.2 \frac{\rho_R}{\rho}, \quad (18b)$$

$$Q_D \equiv \frac{\sqrt{D}\Omega^{3/2}}{\pi G \Sigma}; \quad (18c)$$

which incorporate the effective pressure (with T for Toomre), layer thickness (with R for Roche) and turbulent diffusion of solids (D), respectively. We express Q_R in terms of the critical density of equation (3) and the midplane particle density $\rho = \Sigma/(\sqrt{2\pi}h)$.

The three Q values are not independent. We relate them to the level of turbulent stirring in §3.2, which has the added benefit of reducing the number of input parameters to two. This approach allows the complete survey of solutions in §5. First we describe some general properties of the dispersion relation.

3.1.2. Standard GI: No Drag or Diffusion

We recover standard axisymmetric GI by ignoring gas drag, $\tau_s \rightarrow \infty$, and turbulent diffusion, $Q_D = 0$. The dispersion relation takes the simple form

$$\gamma(\gamma^2 + \mathcal{F}_W) = 0, \quad (19)$$

giving a neutral mode (NM) with $\gamma = 0$ and a pair of density waves (DW) with $\gamma = \pm\sqrt{-\mathcal{F}_W}$. In terms of the dimensional frequency $\omega = i s$, the DW solution obeys the familiar dispersion relation:

$$\omega^2 = \Omega^2 - 2\pi G \Sigma k \cdot \mathcal{T}(kh) + k^2 c^2 \quad (20)$$

The DW solutions correspond to traveling waves if $\mathcal{F}_W > 0$, and gravitational collapse for $\mathcal{F}_W < 0$. In the thin disk limit where $\mathcal{T} \rightarrow 1$, we recover the Toomre (1964) instability criterion, $Q_T < Q_{T,\text{crit}} = 1$, with the marginally unstable wavenumber, $\mathcal{K}_{\text{crit}} = 1$. In physical units $\lambda = \lambda_G$.

We now consider the effects of our finite thickness correction. First note that \mathcal{T} enforces a Roche-like limit. Even for a perfectly cold layer with $Q_T = 0$, instability

($\mathcal{F}_W < 0$) requires $Q_R \leq 2$. If we fix $Q_T = Q_R$ (the standard $c = \Omega h$ relation from Keplerian orbital oscillations or hydrostatic balance) then $Q_{T,\text{crit}} \approx 0.55$ and $\mathcal{K}_{\text{crit}} \approx 1.2$. This solution ignores the compression of the layer by self-gravity which enhances the vertical gravity by a factor

$$\psi \equiv 1 + \frac{2\pi G \Sigma}{\Omega^2 h} = 1 + \frac{2}{Q_R}. \quad (21)$$

Vertical hydrostatic balance then gives a relation $Q_T = Q_R \sqrt{1 + 2/Q_R}$ (as in Ward 2000). In this case $Q_{T,\text{crit}} \approx 0.77$ with $Q_{R,\text{crit}} \approx 0.26$ and $\mathcal{K}_{\text{crit}} \approx 1.05$, in good agreement with more rigorous analyses (Goldreich & Lynden-Bell 1965). Thus the finite thickness correction to Toomre instability is relatively modest when vertical compression is included. Secular GI has much longer wavelengths, making the finite thickness corrections insignificant in most cases of interest.

3.1.3. Proof of Generic Instability

We now prove that the full dispersion relation of equation (16) always yields a growing mode for some wavelengths, for *any* choice of stability parameters or stopping time. This result is fundamentally different from standard GI, as we just reviewed. The existence of a growing mode does not assure its astrophysical relevance, so we impose extra constraints in §4.1. Nevertheless, the following proof demonstrates that instability is quite robust and does not depend, e.g. on the specific turbulence model we adopt in §3.2. Moreover we know that a numerical root finding algorithm has failed if it does not identify a growing mode.

Since the dispersion relation is a cubic in γ with real coefficients, we can use Descartes' "rule of signs" to show that there is always at least one real, positive root (and thus a growing mode). We write equation (16) as

$$\gamma^3 + C_2 \gamma^2 + C_1 \gamma + C_0 = 0 \quad (22)$$

with

$$C_2 = 2/\tau_s + \mathcal{K}^2 Q_D^2 > 0$$

$$C_1 = \mathcal{F}_W + \tau_s^{-2} + 2\mathcal{K}^2 Q_D^2 / \tau_s$$

$$C_0 = (\mathcal{F}_W - 1)/\tau_s + C' \mathcal{K}^2 / \tau_s$$

$$C' = Q_D^2 (\tau_s + \tau_s^{-1}) > 0$$

We use the long wave limit where $\mathcal{F}_W \rightarrow 1 - 2\mathcal{K}$ as $\mathcal{K} \rightarrow 0$. We have $C_2 > 0$ and $C_1 > 0$ since $\mathcal{F}_W > 0$ for long waves. The C_0 coefficient is negative if $C' \mathcal{K}^2 < 1 - \mathcal{F}_W$, which holds because $C' \mathcal{K}^2 < 2\mathcal{K}$ as $\mathcal{K} \rightarrow 0$. Thus we can ensure (for some \mathcal{K}) one, and only one, sign change in the coefficients, which guarantees a positive root. Q.E.D.

3.2. Turbulence Model

We will now describe the response of solid particles to turbulent stirring. This section culminates in equation (33), which expresses the self-gravity parameters Q_T , Q_R and Q_D in terms of the stopping time τ_s and a parameter

$$Q_\alpha = \frac{\sqrt{D_g} \Omega^{3/2}}{\pi G \Sigma} = \sqrt{\alpha} \frac{c_g \Omega}{\pi G \Sigma} \quad (23)$$

that compares turbulent stirring to particle self-gravity. Here D_g is the strength of turbulent diffusion in the gas, and

$$\alpha = D_g \Omega / c_g^2 \quad (24)$$

is a more familiar dimensionless measure of turbulent intensity with c_g the sound speed in the gas. We discuss in §7 why this diffusive α can be much smaller than the (purposefully similar, but not equivalent) parameter used in studies of accretion disks.

This subsection makes heavy use of the results of Youdin & Lithwick (2007, hereafter YL07). We extend their results in our equations (29)–(31) by including self-gravity in the scale height calculation.

3.2.1. Turbulence and α Values

We adopt a relatively simple model for disk turbulence. As in Kolmogorov turbulence, the largest eddies dominate the energy budget. We further assume that $t_{\text{eddy}} = 1/\Omega$, i.e. the turnover time of these integral scale eddies equals the orbital time scale. This assumption requires a turbulent speed $V_{\text{eddy}} = \sqrt{\alpha} c_g$ and eddy length $\ell_{\text{eddy}} = \sqrt{\alpha} c_g / \Omega$ so that $t_{\text{eddy}} = \ell_{\text{eddy}} / V_{\text{eddy}}$ and the gas diffusivity $D_g = V_{\text{eddy}} \ell_{\text{eddy}}$ satisfies equation (24), as noted by Cuzzi et al. (2001).

We further assume that the turbulent kinetic energy is isotropic with partial correlation between the radial, u_g , and azimuthal, v_g , turbulent speeds. Specifically, we take

$$\langle u_g^2 \rangle = \langle v_g^2 \rangle = \langle w_g^2 \rangle = 4 \langle u_g v_g \rangle \equiv V_{\text{eddy}}^2 \equiv \alpha c_g^2 \quad (25)$$

where w_g is the vertical turbulent speed and angular brackets indicate time-averages. While the velocity correlations are of modest significance for particle stirring (see YL07) they are a primary concern for the angular momentum transport and the accretion of gas onto the protostar, see §7.

3.2.2. Particle Response to Turbulence

The radial particle diffusion coefficient is given by equation (36) of YL07. With our equation (25)

$$D = \frac{1 + \tau_s + 4\tau_s^2}{(1 + \tau_s)^2} D_g. \quad (26)$$

The detailed τ_s dependence will vary with the exact properties of the turbulent cascade. The important result is that $\tau_s \ll 1$ solids diffuse at a rate similar to any passive contaminant, $D \simeq D_g$. Particle clumping in turbulence could lower D , but we do not have a good model for this effect. As particles decouple and $\tau_s \gg 1$, diffusivity drops as $D/D_g \propto \tau_s^{-2}$. Thus a rough value of the so called Schmidt number is $\text{Sc} \equiv D_g/D \sim 1 + \tau_s^2/4$. This result differs from the often used, but incorrect, value of $\text{Sc}_{\text{CDC}} = 1 + \tau_s$ in Cuzzi et al. (1993), see YL07.

Since we are interested in radial pressure support, the relevant particle dispersion speed, c , is the radial component. We use

$$c = \frac{\sqrt{1 + 2\tau_s^2 + (5/4)\tau_s^3}}{1 + \tau_s^2} \sqrt{\alpha} c_g. \quad (27)$$

A similar result for c follows from equation (33a) of YL07 and our equation (25). In equation (27) we adopt a simpler τ_s dependence. This choice is purely for convenience

and gives indistinguishable numerical results. These details are not crucial. Völk et al. (1980) showed that the same basic scaling, $c \sim V_{\text{eddy}}/(1 + \tau_s)^{1/2}$, holds when the orbital dynamics included by YL07 are ignored.

The particle scale height, h , usually obeys the well known result

$$h = h_{\text{std}} \equiv \sqrt{\frac{\alpha}{\tau_s}} h_g. \quad (28)$$

where $h_g = c_g/\Omega$ is the hydrostatic gas scale height (for weak self-gravity). Intuitively the particle layer narrows for weaker turbulence or more loosely coupled solids. Cuzzi et al. (1993) derive the result for $\tau_s \ll 1$ by comparing settling and diffusive time scales. Carballido et al. (2006) show that the same scaling also holds for $\tau_s \gg 1$ using both simulations and velocity diffusion arguments.

Corrections to h_{std} arise from self-gravity and the perfect mixing limit. Including self-gravity gives

$$h = h_\psi = \sqrt{\frac{\alpha}{\tau_s \psi}} h_g \quad (29)$$

with the amplification of vertical gravity by ψ from equation (21). To derive this result (which as far as we know is new) for $\tau_s \ll 1$, equate the shortened settling time $t_{\text{sett}} = 1/(\Omega \tau_s \psi)$ with the diffusion time $t_{\text{diff}} = h^2/D_g$. The possibility that turbulent intensity varies with self-gravity is not considered here, but could be included to model gravito-turbulence. The same expression holds for $\tau_s \ll 1$, where the vertical oscillation frequency shortens to $\omega_z = \Omega/\sqrt{\psi}$. The random velocity, $c_z = \sqrt{\alpha/\tau_s} c_g$, imparted by turbulent fluctuations gives $h_\psi = c_z/\omega_z$, which obeys equation (29). Since ψ itself depends on h , the explicit solution reads

$$h_\psi = \sqrt{\ell_G^2 + \frac{\alpha}{\tau_s} h_g^2} - \ell_G \quad (30)$$

which reduces to equation (28) for small ℓ_G , which is defined in equation (13).

Perfect mixing with the gas imposes an upper limit $h \leq h_g$. Following YL07 [their equation (28)] we enforce this limit to give our most general result

$$h = h_{\text{full}} = \sqrt{\frac{\alpha}{\tau_s \psi + \alpha}} h_g. \quad (31)$$

Again there is an implicit h dependence in ψ .

The ratio of particle to gas density, important for drag feedback effects, is

$$\frac{\rho}{\rho_g} = \frac{\Sigma}{\Sigma_g} \frac{h_g}{h} \approx Z \sqrt{\frac{\tau_s}{\alpha}} \quad (32)$$

which uses h_{std} for illustration, though equation (31) is used in numerical calculations.

3.2.3. Dimensionless Stability Parameters

We can now express the Q parameters from equation (18) in terms of Q_α and τ_s . Applying equations (26), (27) and (31) gives

$$Q_T = \frac{\sqrt{1 + 2\tau_s^2 + (5/4)\tau_s^3}}{1 + \tau_s^2} Q_\alpha, \quad (33a)$$

$$Q_R = \frac{\sqrt{1 + Q_\alpha^2/\tau_s(1 + \alpha/\tau_s)} - 1}{1 + \alpha/\tau_s}, \quad (33b)$$

$$Q_D = \frac{\sqrt{1 + \tau_s + 4\tau_s^2}}{1 + \tau_s^2} Q_\alpha. \quad (33c)$$

Note that equation (33b) depends on α in addition to Q_α and τ_s . To reduce to only two parameters, we take the $\alpha \ll \tau_s$ limit, equivalent to using equation (29), and obtain

$$Q_R \approx \sqrt{1 + Q_\alpha^2/\tau_s} - 1. \quad (34)$$

This approximation is acceptable on a few counts. First it is conservative, since we are removing the restriction that $h < h_g$ and allowing artificially low particle densities. Second, relevant growth rates typically have $\alpha < \tau_s$, as seen in Figs. 1 and 4. Finally, even for $\alpha > \tau_s$, wavelengths are long enough that the finite thickness correction is negligible, $kh = \mathcal{K}Q_R \ll 1$, even with the artificially enhanced h . Thus equation (34) is always acceptable for our study, even though equation (33b) is more general.

For $\tau_s \ll 1$, $Q_T \simeq Q_D \simeq Q_\alpha$ is particularly simple. It reflects the result that tightly coupled particles share the same velocity and diffusion coefficient as the dominant turbulent eddies. Furthermore $Q_R \simeq Q_\alpha/\sqrt{\tau_s} \gg 1$ for weak self-gravity, reflecting the fact that $h \gg c\Omega$ for turbulently-stirred small particles.

3.3. Disk Model

3.3.1. A Scalable Minimum Mass Nebula

For numerical evaluations we adopt a reference minimum mass Solar Nebula model, with scaling factors to explore the effect of changing disk parameters. In our formulation (similar to CY10 which has further discussion) the surface density of solids and gas and the midplane temperature are

$$\Sigma = 20 Z_\% F R_{\text{AU}}^{-3/2} \text{ g/cm}^2, \quad (35a)$$

$$\Sigma_g = 2000 F R_{\text{AU}}^{-3/2} \text{ g/cm}^2, \quad (35b)$$

$$T = 200 f_T R_{\text{AU}}^{-1/2} \text{ K}, \quad (35c)$$

where $R_{\text{AU}} \equiv R/\text{AU}$ is the disk radius in AU. The “metallicity” $Z \equiv \Sigma/\Sigma_g \equiv 0.01 Z_\%$ and the mass normalization F are two important adjustable parameters. Water and methane condensation at “ice lines” can give sharp transitions in Σ . We do not model this in detail, because a range of enrichment processes also affect Z .

We adopt a cool disk temperature characteristic of passive flared disks around Sun-like stars, and fix $f_T = 1$ in this work. Our profile is modestly steeper than the $T \propto R^{-3/7}$ of (Chiang & Goldreich 1997), just to give analytic scalings that are both simpler and similar to the classic Hayashi (1981) model. This idealization has little effect, given the weak dependence of our results on such small temperature changes.

The disk mass within a cutoff radius R_{out} is

$$M_{\text{disk}} = 0.02 F(1 + Z) \sqrt{R_{\text{out}}/(50 \text{ AU})} M_\odot. \quad (36)$$

The gas disk is gravitationally stable since its Toomre parameter,

$$Q_g \equiv \frac{c_g \Omega}{\pi G \Sigma_g} \approx 40 \frac{\sqrt{f_T m_*}}{F} R_{\text{AU}}^{-1/4}, \quad (37)$$

satisfies $Q_g > 1$, with m_* the stellar mass in solar masses. Our measure of particle self-gravity relative to turbulent stirring, from equation (23),

$$Q_\alpha = \sqrt{\alpha} \frac{Q_g}{Z} \approx 1.3 \sqrt{\frac{\alpha}{10^{-7}}} \frac{\sqrt{f_T m_*}}{F Z_\%} R_{\text{AU}}^{-1/4}, \quad (38)$$

exceeds unity in most disk models.

The midplane gas density is

$$\rho_g = \frac{\Sigma_g}{\sqrt{2\pi} h_g} \approx 2 \times 10^{-9} F \sqrt{\frac{m_*}{f_T}} R_{\text{AU}}^{-11/4} \frac{\text{g}}{\text{cm}^3} \quad (39)$$

for a vertically isothermal disk, ignoring compression from the weight or self-gravity of the solids (Nakagawa et al. 1986; Wakita & Sekiya 2008).

The mean free path for molecular collisions,

$$\ell_g \approx 1 \sqrt{\frac{f_T}{m_*}} \frac{1}{F} R_{\text{AU}}^{11/4} \text{ cm}, \quad (40)$$

is inversely proportional to ρ_g and important for the drag forces that we now consider.

3.3.2. Particle Properties and Gas Drag

Our model considers a population of particles with uniform radius, a , and internal density, $\rho_\bullet = 2 \text{ g/cm}^3$. The general dependence of drag forces on particle, gas and flow properties is complex, but approximate formulae apply to distinct parameter regimes (Adachi et al. 1976; Weidenschilling 1977). The stopping time $t_{\text{stop}} \equiv \Delta V/f_{\text{drag}}$ is the ratio of the relative flow speed to the drag acceleration. Fig. 1 plots the dimensionless $\tau_s \equiv \Omega t_{\text{stop}}$ for particle sizes from $a = 1 \text{ mm}$ to 30 cm for both a minimum mass disk with $F = 1$ and a gas depleted disk with $F = 0.1$.

Epstein drag is the most relevant regime. It applies for small sizes, $a \lesssim (9/4)\ell_g$, and low gas densities, with

$$\tau_s = \tau_{\text{Ep}} = \frac{\rho_\bullet a}{\sqrt{2\pi} \Sigma_g} \approx 2.5 \times 10^{-4} \frac{a_{\text{mm}}}{F} R_{\text{AU}}^{3/2}, \quad (41)$$

where $a_{\text{mm}} = a/\text{mm}$. Epstein drag is readily identifiable in Fig. 1 as the region where τ_s increases with R (outside 0.3 AU for 1 mm and outside $\sim 2 \text{ AU}$ for 30 cm).

Stokes drag applies to particles too large for Epstein drag, and a Reynolds number, $\text{Re} = a\Delta V/(c_g \ell_g) \lesssim 1$. The stopping time for viscous Stokes drag

$$\tau_s = \tau_{\text{Sto}} = \tau_{\text{Ep}} \frac{4a}{9\ell_g} \approx 1 \times 10^{-3} a_{\text{mm}}^2 \sqrt{\frac{m_*}{f_T}} R_{\text{AU}}^{-5/4}, \quad (42)$$

applies in the inner disk, where gas densities are high, and for small solids that do not experience a large headwind. In Fig. 1 we identify the Stokes regime for 1 mm and 3 cm solids where τ_s declines sharply with radius. The brief overlap of the $F = 1.0$ and $F = 0.1$ (gas depleted) curves demonstrates that τ_{Sto} does not depend on gas density.

In Fig. 1, turbulent drag applies in regions where τ_s has a relatively flat dependence on R . Appendix A discusses non-linear drag regimes. Unless stated otherwise, the numerical estimates and scalings in this work use Epstein drag.

3.3.3. Radial Drift

Particles experience a net inward drift because they encounter a headwind in pressure-supported gas disks. The orbital speed of a purely gas disk is slower than Keplerian by a speed $\eta\Omega R$, where

$$\eta \equiv -\frac{\partial P/\partial R}{2\rho_g\Omega^2 R} \approx 1.3 \times 10^{-3} \frac{f_T}{m_*} \sqrt{R_{\text{AU}}} \quad (43)$$

measures the radial gradient of the gas pressure, P .

The radial drift time scale, defined as the time to reach the star if the speed were to remain constant, is

$$t_{\text{drift}} \equiv \frac{R}{|v_{\text{drift}}|} = f_{\text{in}} \frac{1 + \tau_s^2/f_{\text{in}}}{2\eta\tau_s\Omega}. \quad (44)$$

This equation includes the inertial correction factor (Nakagawa et al. 1986; Youdin & Chiang 2004)

$$f_{\text{in}} \equiv (\rho_{\text{tot}}/\rho_g)^2 = (1 + \rho/\rho_g)^2, \quad (45)$$

which is unity in the test particle limit that, $\rho \ll \rho_g$. Test particles with $\tau_s = 1$ have the shortest drift time

$$\min(t_{\text{drift}}) = (\eta\Omega)^{-1} \approx 120\sqrt{m_*}R_{\text{AU}}/f_T \text{ yr}. \quad (46)$$

The “meter-size barrier” refers roughly to the size of $\tau_s = 1$ bodies in the inner disk. As Fig. 1 shows, \sim cm-sized bodies drift fastest in the outer disk. Particle inertia both lengthens the shortest drift time — as $\min(t_{\text{drift}}) \propto \sqrt{f_{\text{in}}} = \rho_{\text{tot}}/\rho_g$ — and shifts it to larger $\tau_s = \rho_{\text{tot}}/\rho_g$.

In the small particle regime, with $\tau_s < 1$

$$t_{\text{dr,tight}} \approx \frac{f_{\text{in}}}{2\eta\Omega\tau_s} \approx 2.5 \times 10^5 \frac{f_{\text{in}}F\sqrt{m_*}}{a_{\text{mm}}f_T\sqrt{R_{\text{AU}}}} \text{ yr}. \quad (47)$$

Larger particles drift faster in the test particle limit. However inertia can cancel this trend. With a fixed level of weak turbulence that ensures $\rho \gg \rho_g$, equation (32) gives $f_{\text{in}} \approx Z^2\tau_s/\alpha$, canceling the explicit size dependence above.

For larger particles with $\tau_s \gtrsim 1$, the drift time increases with size and radial distance as

$$t_{\text{dr,loose}} \approx \frac{\tau_s}{2\eta\Omega} \approx 2.7 \times 10^5 \frac{a_m\sqrt{m_*}}{f_T F} R_{50}^{5/2} \text{ yr}, \quad (48)$$

where $R_{50} \equiv R/(50 \text{ AU})$ and $a_m = a/(1 \text{ m})$. There is no inertial correction for large solids with $\tau_s \gg \rho_{\text{tot}}/\rho_g$, because drag forces are too weak to slow the headwind.

4. NUMERICAL RESULTS

We now present the main numerical results for the gravitational collapse of solids in a turbulent gas disk. The fiducial protoplanetary disk models of §3.3 provide the physical values. Our most commonly plotted quantities are the growth time, wavelength and mass of solids in a collapsing ring, which we define as

$$t_{\text{grow}} \equiv [\Omega \text{Real}(\gamma)]^{-1}, \quad (49a)$$

$$\lambda \equiv \lambda_G/\mathcal{K}, \quad (49b)$$

$$M_{\text{ring}} \equiv 2\pi\lambda R\Sigma. \quad (49c)$$

To obtain these quantities, we solve the dispersion relation in equation (16) for the fastest growing modes. This solution involves maximizing the real part of the

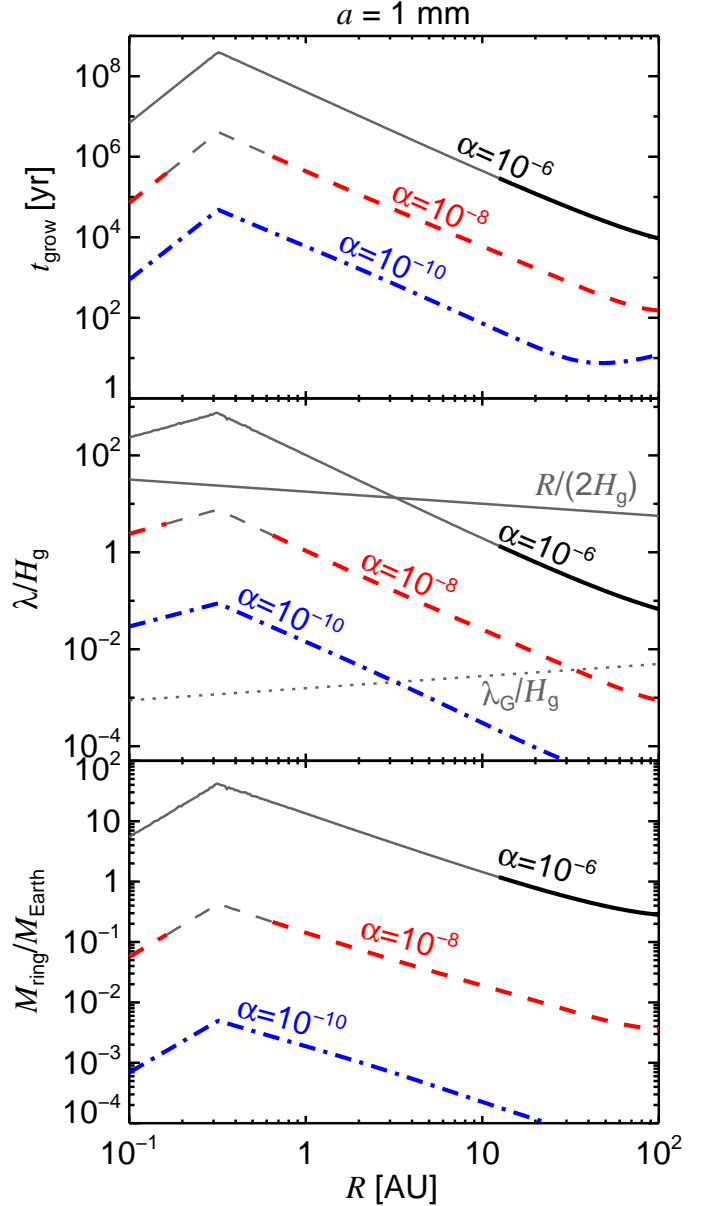


Figure 2. Properties of the dissipative gravitational collapse of mm-sized particles vs. disk radius in the reference model for several values of the turbulent diffusion parameter $\alpha = 10^{-6}, 10^{-8}, 10^{-10}$ (solid, dashed, and dot-dashed curves). Grey sections of curves denotes when growth is too slow. While remarkably low, these α values do not preclude stronger turbulence away from the mid-plane, and also ignore the possibility that small-scale clumping reduces diffusion. *Top:* Growth times in years. *Middle:* Wavelengths in AU compared to (half) the disk radius, the gas scale height h_g , and the standard GI wavelength λ_G (brown, black and grey dotted lines, respectively). *Bottom panel:* Mass of collapsing ring in Earth masses. Subsequent fragmentation yields much smaller planetesimal masses. See text (§4.2) for details.

(dimensionless) growth rate γ with respect to the (dimensionless) wavenumber \mathcal{K} , see §5 for details. We quote a ring mass because the modes are axisymmetric, but emphasize that this is *not* a predicted planetesimal mass. Gravitational fragmentation during later stages of collapse would likely produce smaller bodies.

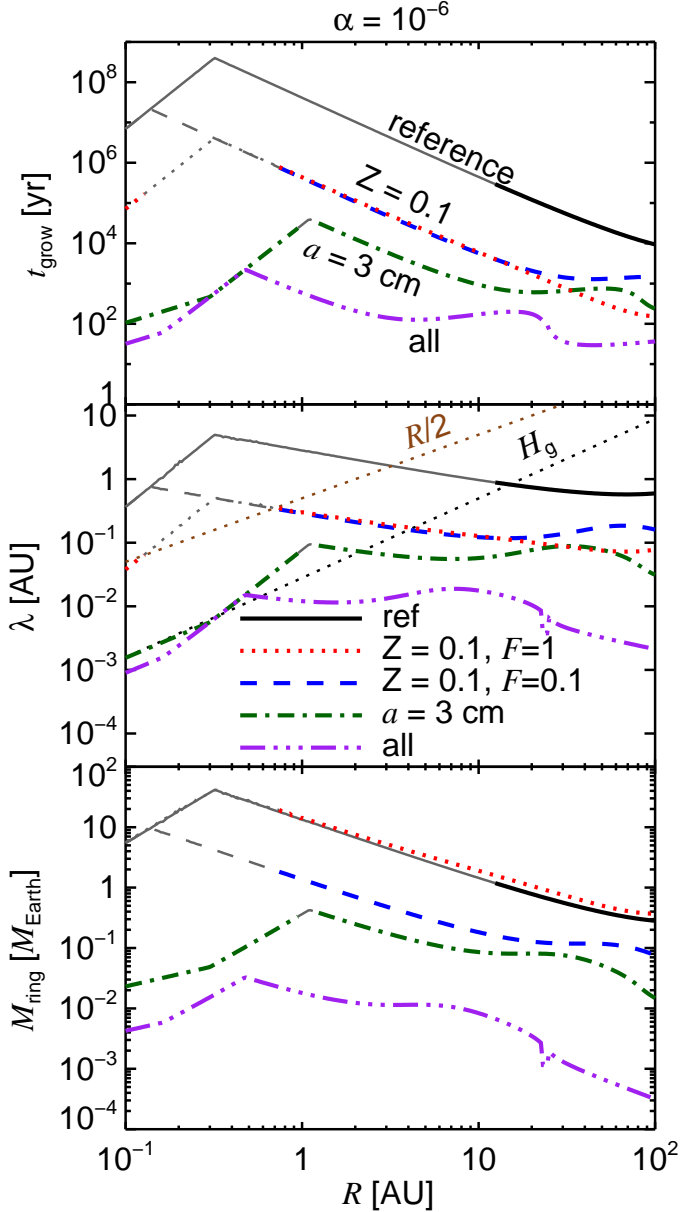


Figure 3. Similar to Fig. 2 with turbulence held fixed at $\alpha = 10^{-6}$ and other parameters varied. The reference model with $a = 1$ mm solids (solid black curve) is shown again for comparison. Larger $a = 3$ cm solids (green dot-dashed curves) give faster and smaller scale growth. Increasing the particle-to-gas surface density ratio to $Z = 0.1$ also promotes growth, whether the gas content stays fixed ($F = 1.0$, red dotted curve) or is depleted ($F = 0.1$, blue dashed curve). With both gas depletion and 3 cm particles (purple triple-dot-dashed curve) growth is even faster and smaller scale. The kink in λ and M_{ring} at 20 AU is a real feature (not a numerical artifact) as explained in §5.4.

4.1. When is Collapse Relevant?

Since our instability always produces growth (see §3.1.3), we require the following conditions be met for collapse to be relevant:

$$t_{\text{grow}} < t_{\text{drift}}, \quad (50a)$$

$$t_{\text{grow}} < t_{\text{life}} = 10^6 \text{ yrs}, \quad (50b)$$

$$\lambda < R/2. \quad (50c)$$

Growth must be faster than both the radial drift time of equation (44) and an assumed 10^6 year disk lifetime. Also the wavelength must fit comfortably inside the disk. The drift requirement is usually the most stringent.

The characteristic growth time, t_{grow} , only represents a single e-folding. Many e-foldings are generally required for significant density growth. The actual time to form planetesimals depends on seed perturbation amplitudes and non-linear evolution, both beyond this analysis. To avoid introducing another uncertain parameter, we simply comment that many e-foldings are possible because growth could proceed over several t_{drift} , especially since the growth of particle overdensities will slow the actual drift timescales. Moreover disk lifetimes could extend to ~ 10 Myr, though this is less often the relevant constraint. These issues touch on the application of a local instability to a globally evolving disk, as discussed further in §6.3.

4.2. Growth for Fixed α

Our reference model consists of particles with radius $a = 1$ mm in our minimum mass disk model with $F = 1$ and $Z = 0.01$. Fig. 2 plots growth times and wavelengths for different levels of the turbulent parameter α with all other disk quantities fixed. Curves become light grey whenever the conditions of equation (50) are not met. Low α values are required for collapse in our nominal disk. With $\alpha = 10^{-6}$ collapse only proceeds outside 10 AU. For $\alpha = 10^{-8}$ collapse occurs everywhere except a region near $R = 0.3$ AU, where the peak in growth times corresponds with the minima in stopping times at the Epstein-Stokes transition, see Fig. 1.

Increasing the strength of turbulence gives slower collapse that acts over long wavelengths. Our physical derivation of equations (9) and (10) explains the basic scalings. Including the more precise coefficients of equation (56), the growth time for $\tau_s \ll 1$ becomes

$$t_{\text{grow}} \simeq \frac{1}{\Omega} \frac{\alpha Q_g^2}{Z^2 \tau_s^2}. \quad (51)$$

The more rapid collapse in the outer disk occurs mainly due to weaker drag coupling, with larger τ_s . This aerodynamic effect overcomes the radial increase in orbital time scales. Stronger self-gravity, as measured by Q_g , also assists growth in the outer disk.

Substituting numerical scalings for Epstein drag and the reference disk, the growth properties are

$$\frac{t_{\text{grow}}}{\text{yr}} \approx 4 \times 10^5 \frac{\alpha_{-8} f_T \sqrt{m_*}}{Z_{\%}^2 a_{\text{mm}}^2} R_{\text{AU}}^{-2}, \quad (52a)$$

$$\frac{\lambda}{\text{AU}} \approx 0.028 \frac{\alpha_{-8} f_T}{Z_{\%} a_{\text{mm}}} R_{\text{AU}}^{-1/2}, \quad (52b)$$

$$\frac{M_{\text{ring}}}{M_{\oplus}} \approx 0.13 \frac{\alpha_{-8} F f_T}{a_{\text{mm}} R_{\text{AU}}}, \quad (52c)$$

where turbulence is normalized to $\alpha_{-8} = \alpha/10^{-8}$. Different scalings (not given) apply to the Stokes regime at small R .

Significantly faster and smaller scale growth is possible with deviations from the reference model. Fig. 3 explores the effect of varying disk parameters while keeping $\alpha = 10^{-6}$ fixed. Larger particle radii with bigger τ_s give

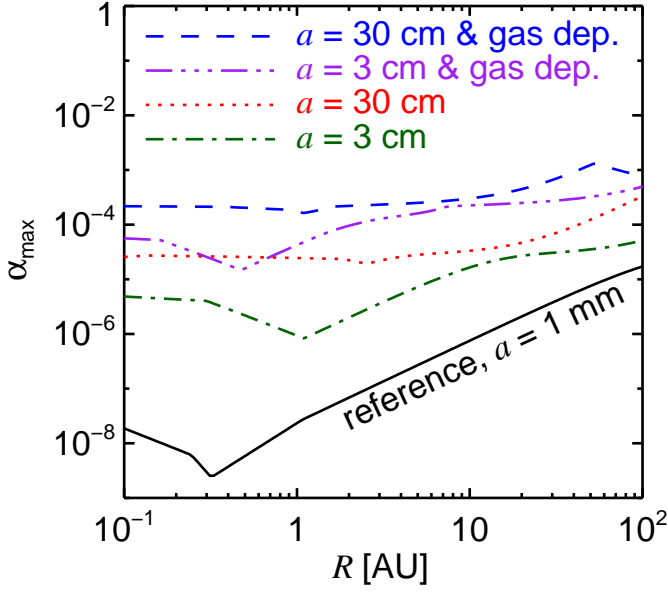


Figure 4. The maximum strength of turbulence, given by α_{\max} , which allows collapse. The criteria for sufficiently rapid collapse are described in equation (50). The α_{\max} values increase with particle size as shown for $a = 1$ mm (black solid curve), $a = 3$ cm (dash-dotted green curve) and $a = 30$ cm (dotted red curve). The enrichment of the solid to gas ratio also increases α_{\max} . We show this for gas depleted models with $Z = 0.1$ and $F = 0.1$ for both $a = 3$ cm (triple-dot dashed purple curve) and $a = 30$ cm (dashed blue curve). See Fig. 5 for the properties of modes with $\alpha = \alpha_{\max}$.

much faster growth, as the $a = 3$ cm (dot-dashed green) curves demonstrate. Deviations from powerlaw behavior at large R arise from the transition to loose coupling, $\tau_s \gtrsim 1$, which §5 explores in more detail.

Higher particle fractions, Z , produce faster growth by increasing particle self-gravity. As an added benefit, more particle inertia slows radial drift. Whether Z increases by adding solids or depleting gas (or some combination) the effect on growth rates and wavelengths is similar. Indeed equations (52a) and (52b) are independent of F . While a higher mass disk has stronger particle self-gravity (at fixed Z), higher gas densities cancel this effect by giving tighter coupling (in the Epstein, but not Stokes, regime). The ring mass in equation (52c) does increase with F . At fixed Z , there is obviously more particle mass in a fixed area when the total mass rises.

Fig. 3 demonstrates these metallicity effects for two high $Z = 0.1$ models, with $F = 1$ (red dotted curve) and $F = 0.1$ (blue dashed curve). The overlap in t_{grow} and λ is evident at intermediate R . The curves separate at small R with the transition to Stokes drag, and at large R with the transition to $\tau_s > 1$, which occurs at smaller R in the gas depleted model.

The effect of particle growth and enhanced metallicity can interact constructively. The purple (triple dot-dashed) curve labelled “all” shows this result by combining the $a = 3$ cm, $Z = 0.1$ and $F = 0.1$ parameters. Growth times are even faster and ring masses are much smaller when these effects are considered together.

4.3. Maximum Allowed Turbulence

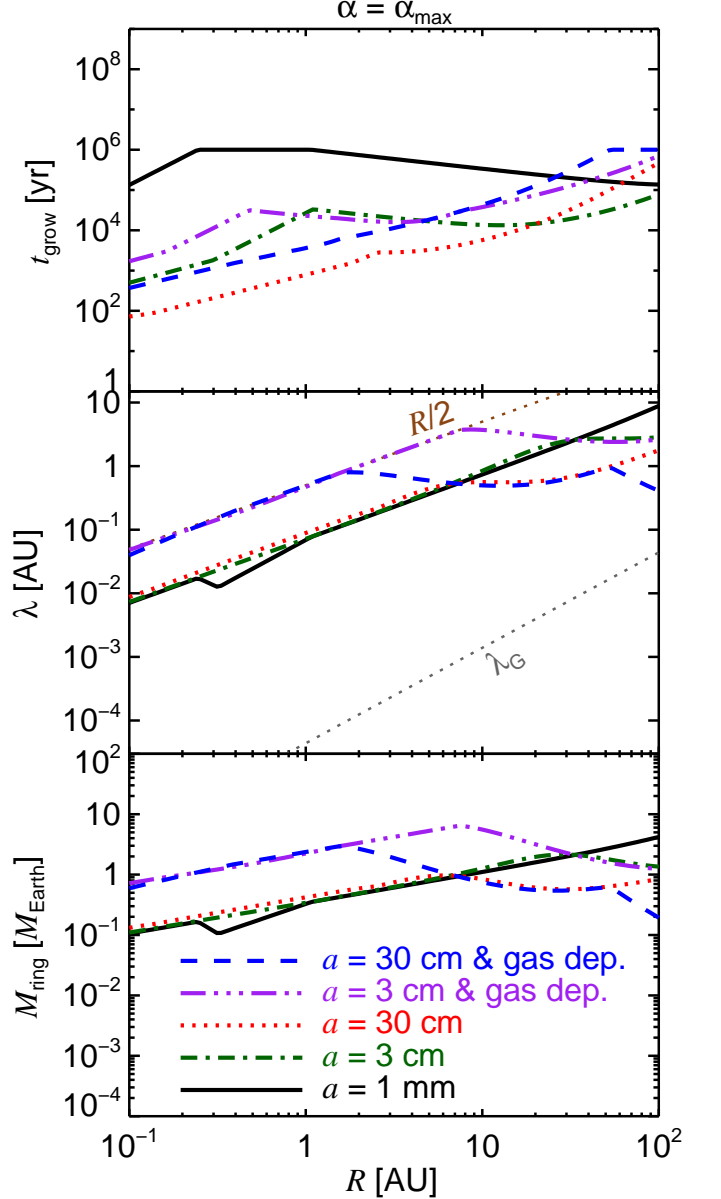


Figure 5. Properties of modes with the maximum allowed turbulence, $\alpha = \alpha_{\max}$, for the same models as in Fig. 4. We can visually identify which constraint sets α_{\max} : if $t_{\text{grow}} = 10^6$ yr, it is the disk lifetime, and when $\lambda = R/2$, it is the disk size. Otherwise the drift time scale is the relevant constraint. See the text (§4.3) for an explanation of the overlapping wavelengths and masses.

Instead of fixing the level of turbulence, we now determine the strongest level of turbulence that allows collapse. Fig. 4 plots α_{\max} , the largest α value for which growing modes satisfy all the constraints of equation (50). Not surprisingly, the same factors that give faster growth — larger solids and higher particle mass fractions — also yield larger α_{\max} .

The constraints on turbulence are strongest inside a few AU, where $\alpha \lesssim 10^{-4}$ is required for sizes up to 30 cm, even with some metallicity enrichment. While the outer disk generally allows collapse with stronger turbulence, it still requires $\alpha \lesssim 10^{-3}$.

Analytic fits to α_{\max} come with even more caveats

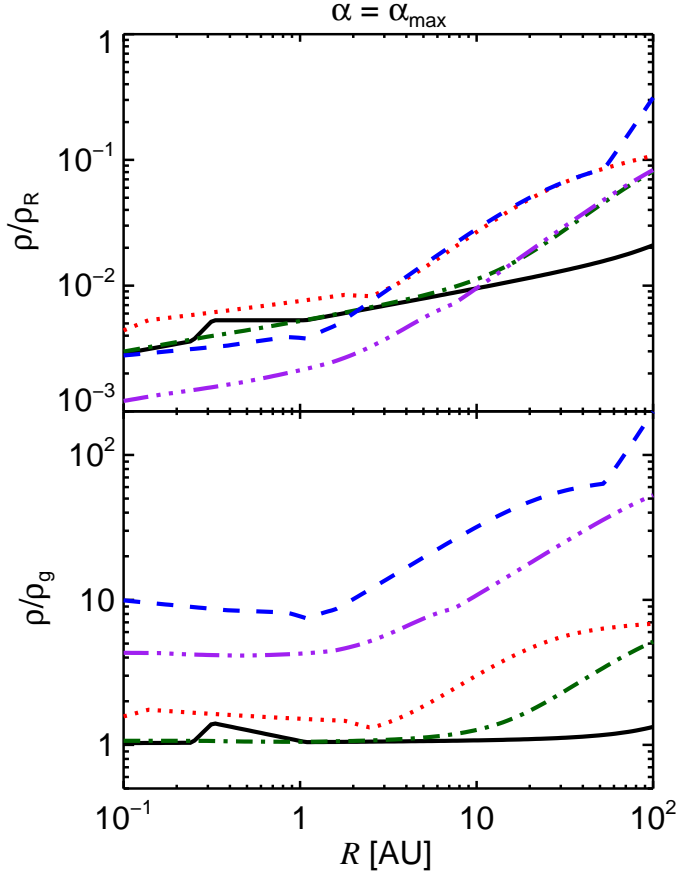


Figure 6. The minimum particle density, ρ , that gives collapse, obtained by setting $\alpha = \alpha_{\max}$, using the same disk models as in Figs. 4 and 5. (*Top*): ρ plotted relative to the Roche density [equation (3)], demonstrating that secular GI occurs well below this standard threshold. (*Bottom*): ρ plotted relative to gas density to show that drag feedback is always significant.

than usual, but are useful for a rough understanding. When drift speeds are the limiting constraint, the condition $t_{\text{grow}} < t_{\text{drift}}$ sets

$$\alpha_{\max} \simeq \frac{\tau_s Z^2 f_{\text{in}}}{2\eta Q_g^2} \approx 0.6 \times 10^{-8} \frac{a_{\text{mm}} Z_{\%}^2 F f_{\text{in}}}{f_T^2} R_{\text{AU}}^{3/2}, \quad (53)$$

for $\tau_s \ll 1$. (Again the numerical scalings apply only to Epstein drag.) This basic scaling explains many of the curves in Fig. 4, e.g. the $a = 1$ mm curve beyond 1 AU and segments of the $a = 3$ cm curves. For transitional $\tau_s \sim 1$ values, α_{\max} has a flat R dependence, e.g. for the $a = 3$ cm curve beyond 10 AU. For $\tau_s \gg 1$, a similar constraint to equation (53) holds, again with $\alpha_{\max} \propto R^{3/2}$ as seen in the $a = 30$ cm curve beyond 20 AU. The kink near 60 AU in the gas depleted $a = 30$ cm curve occurs because the disk lifetime constraint becomes relevant.

The upper limit to allowed turbulence, α_{\max} , varies greatly with disk model and has no absolute upper limit. The range of models we consider all require $\alpha_{\max} \lesssim 10^{-3}$, and lower values in the inner disk.

4.3.1. Mode Properties

To better understand these limits on turbulence, we plot the properties of modes with $\alpha = \alpha_{\max}$ in Fig. 5.

These growth times, wavelengths and masses here take the largest allowed values for a given disk model. For $\alpha < \alpha_{\max}$, collapse is more rapid and smaller scale.

Fig. 5 allows us to readily identify the relevant constraint that sets α_{\max} . The 10^6 year lifetime constraint applies when the drift time scale is even longer. As shown in the top panel, this occurs both for loose coupling (beyond 60 AU in the gas depleted $a = 30$ cm model) and for very tight coupling (mm-sized solids near ~ 0.3 AU at the Stokes-Epstein transition). In the middle panel, we see that the disk size constrains wavelengths in our enhanced metallicity models, especially in the inner disk. The radial drift time scale is the relevant constraint on α_{\max} wherever $t_{\text{grow}} \neq 10^6$ yr and $\lambda \neq R/2$.

A striking feature of Fig. 5 is the clustering of the wavelength (and also mass) curves for modes for models with very different τ_s . Substitution of $\alpha_{\max} \propto \tau_s$ from equation (53) into the wavelength result of equation (56) gives

$$\lambda(\alpha_{\max}) \approx \frac{f_{\text{in}}}{2\eta} \lambda_G \approx 0.017 \frac{f_{\text{in}} F Z_{\%} R_{\text{AU}}}{f_T} \text{ AU}, \quad (54)$$

which is independent of τ_s . Thus the maximum wavelength imposed by drift is at least $1/(2\eta)$ larger than the standard GI wavelength, and more if the inertial correction factor, f_{in} , is significant. The gas depleted curves in Fig. 5 have larger maximum wavelengths when $\tau_s < 1$. In these cases, higher particle inertia increases f_{in} values, raising $\lambda(\alpha_{\max})$ to the $\lambda = R/2$ limit.

4.3.2. Minimum Particle Density

The volume density of solids, ρ , is both a measure of self-gravity — though an indirect and imprecise one for secular GI — and a measure of feedback effects that give rise to drag instabilities. When $\alpha = \alpha_{\max}$ we find the minimum ρ required for collapse.

The top panel of Fig. 6 plots $\rho/\rho_R = 0.2/Q_R$. Since $\rho \ll \rho_R$, secular GI occurs well below this standard threshold. While §3.1.3 proved that collapse formally occurs for any Q_R and thus ρ , these numerical results demonstrate collapse for small ρ when astrophysical constraints are imposed. Below the ρ values plotted in Fig. 6, collapse is too slow to be relevant.

We emphasize that density is an indirect measure of the relevance of secular GI. Radial diffusion is the dominant stabilizing influence, so the vertical diffusion that sets ρ is more of a side effect. If we consider asymmetric turbulence and vary just the vertical diffusion, ρ would change, but growth rates would be unaffected because the finite thickness correction is negligible for long wavelength modes. Drift rates however would change from the inertial correction.

The bottom panel of Fig. 6 plots the particle density — again, the minimum value required for collapse — relative to the gas density. Remarkably, the minimum mass models with $F = 1.0$ all have $\rho \gtrsim \rho_g$. Applying the drift constraint of equation (53) to the density relation in equation (32) gives

$$\frac{\rho(\alpha_{\max})}{\rho_g} \approx Q_g \sqrt{\frac{2\eta}{f_{\text{in}}}} \approx 2.0 \frac{f_T}{F \sqrt{f_{\text{in}}}}. \quad (55)$$

Thus ρ/ρ_g is unity for $F = 1$ models (when $\tau_s \lesssim 1$) and larger for the gas depleted $F = 0.1$ models, explaining the bottom panel Fig. 6. It is a numerical coincidence

Table 1
Growth across parameter space

| # | Drag | Turbulence | γ | \mathcal{K} |
|-----|----------------|--|---|---|
| I | $\tau_s \ll 1$ | $Q_\alpha \gg \tau_s^{3/2}$ | τ_s^2/Q_α^2 | τ_s/Q_α^2 |
| II | $\tau_s \gg 1$ | $Q_\alpha \gg \sqrt{\tau_s}$ | $0.2/Q_\alpha^2$ | $0.2\tau_s/Q_\alpha^2$ |
| III | $\tau_s \gg 1$ | $\frac{1}{\sqrt{\tau_s}} \ll Q_\alpha \ll \sqrt{\tau_s}$ | $0.7\sqrt{\tau_s}/Q_\alpha$ | $0.5\tau_s/Q_\alpha^2$ |
| IV | $\tau_s \gg 1$ | $Q_\alpha \ll 1/\sqrt{\tau_s}$ | $0.6\left(\frac{\tau_s}{Q_\alpha}\right)^{2/3}$ | $0.4\left(\frac{\tau_s}{Q_\alpha}\right)^{4/3}$ |
| V | $\tau_s \ll 1$ | $Q_\alpha \ll \tau_s^{3/2}$ | $Q_\alpha^{-2/3}$ | $Q_\alpha^{-4/3}$ |

Note. — Col. (1): Region label, see also Fig. 8. Cols. (2) & (3): Region boundaries in τ_s - Q_α space. Cols. (4) & (5): Growth rate and wavenumber of fastest growing modes.

— unrelated to the mechanism of the instability — that standard minimum mass gas disks give $Q_g\sqrt{2\eta}$ of order unity. But this coincidence is significant. The consequence that $\rho \gtrsim \rho_g$ for gravitational collapse means that drag feedback is relevant, and can drive particle-driven turbulence. See §6.3 for more discussion.

5. GROWTH ACROSS PARAMETER SPACE

We now survey the behavior of our instabilities across the full range of parameter space. Varying only two parameters — the stopping time, τ_s , and the strength of turbulence relative to particle self-gravity, Q_α — covers all possibilities, thanks to the turbulence model of §3.2. Once solutions to the dispersion relation of equation (16) are tabulated, translating the conditions in an arbitrary disk model to appropriate values of τ_s and Q_α allows simple look-up of the solutions. Furthermore, we gain both a deeper understanding and a useful check on numerical solutions by understanding the analytic behavior for extreme values of the input parameters.

Fig. 7 maps the growth rate, γ_{fgm} , and wavenumber, \mathcal{K}_{fgm} , of the fastest growing modes versus Q_α and τ_s . We obtain these solutions by maximizing each of the three roots of equation (16) with respect to \mathcal{K} over a well-sampled grid of τ_s and Q_α values.³

Fig. 7 shows the general trend that both growth rates and wavenumbers tend to increase for small Q_α (weaker turbulence or stronger self-gravity) and larger τ_s (bigger solids or weaker drag). However the behavior has a different character in different regions of parameter space. We identify five distinct regions of parameter space, shown in Fig. 8. Table 1 lists the analytic growth rates and wavenumbers by region.

We now describe the behavior by regions, with less relevant cases deferred to appendix B. The transition from secular to standard GI occurs across the region II to III boundary. Fig. 7 shows the corresponding sharp rise in growth rates, explored further in §5.4.

5.1. Region I: Secular, Tight

Region I is the most relevant for planetesimal formation. With $\tau_s \ll 1$ and $Q_\alpha \gg \tau_s^{3/2}$, it describes tight coupling and turbulence that is not incredibly weak.⁴ The

³ For extreme values of the input parameters, renormalizing \mathcal{K} towards a known solution as $\mathcal{K} \rightarrow \mathcal{K}' = \mathcal{K}Q_\alpha^2/\tau_s$ helps avoid numerical difficulties.

⁴ Appendix B.1 shows that the weak turbulence solutions in region V are not relevant.

physical derivation of secular GI in §2 — a simple balance between terminal velocity infall and diffusion — applies in region I.

The fastest growth rate and corresponding wavenumber in region I,

$$\gamma_{\text{I}} = \tau_s^2/Q_\alpha^2, \mathcal{K}_{\text{I}} = \tau_s/Q_\alpha^2, \quad (56)$$

are limiting solutions of the full dispersion relation. They are more exact versions of the physically derived equations (9) and (10). The smooth variation with Q_α , also evident in Fig. 7, demonstrates the insignificance of the $Q_\alpha = Q_{\text{T}} = 1$ transition for $\tau_s \ll 1$.

Growth rates in region I can be as fast as $\gamma \sim 1/\tau_s \gg 1$ at the region V transition. Nevertheless, growth rates are slower than dynamical, $\gamma \ll 1$, because our disk models have $Q_\alpha > 1$ as in equation (38).

The thin disk approximation holds in region I. From equations (34) and (56), $\mathcal{K}_{\text{I}}Q_{\text{R}} \ll 1$ for weak self-gravity with $Q_\alpha > \sqrt{\tau_s}$. The strong self-gravity case of $Q_\alpha < \sqrt{\tau_s} \ll 1$, is less relevant, but still gives only a modest finite thickness correction, $\mathcal{K}_{\text{I}}Q_{\text{R}} \sim 1$. (And thankfully we need not sub-divide region I.)

Density waves are rapidly damped by drag in region I. Growth occurs via the neutral mode, introduced in §3.1.2 as trivial in the absence of drag. The eigenfunctions of the particle response are $\tilde{U}_{\text{I}} = 2\tilde{V}_{\text{I}}/\tau_s = 2\tau_s\tilde{\sigma}_{\text{I}}$. This strong density response, with $\tilde{\sigma} \gg \tilde{U} \gg \tilde{V}$, makes it unlikely that secondary flow instabilities will halt the gravitational collapse once it begins.

Wavenumbers other than the fastest growing can also collapse, at a rate determined by the leading order dispersion relation:

$$\gamma_{\text{I}}(\mathcal{K}) \approx 2\tau_s\mathcal{K} - \mathcal{K}^2Q_\alpha^2. \quad (57)$$

This result has implications for how collapse proceeds for a distribution of particle sizes as discussed in §6.4.

5.2. Region II: Secular, Loose

Region II describes larger, loosely coupled solids with $\tau_s \gg 1$ and relatively weak self-gravity characterized by Toomre's $Q_{\text{T}} = Q_\alpha/\sqrt{\tau_s} \gg 1$. Like the tight coupling case of region I, gravitational collapse involves the secular destabilization of a neutral mode.

The physical derivation of region II growth follows the qualitative description of secular GI in the introduction of Goodman & Pindor (2000). As in §2 we consider a ringlike perturbation of the solids with radial self-gravity $g_r \sim G\sigma$. The orbital speed at the outer (inner) edge of the ring increases (decreases, respectively) by $v \sim g_r/(2\Omega)$ in response. Assuming the gas is unaffected by this acceleration (see §6.5), the extra headwind (tailwind for the inner edge) causes a radial drift towards the ring center. By angular momentum conservation this drift speed is

$$u \sim \frac{2v}{\Omega t_{\text{stop}}} \sim \frac{G\sigma}{\Omega^2 t_{\text{stop}}}. \quad (58)$$

Due to the weaker coupling this scaling differs from equation (4). The collapse rate again follows from mass continuity as

$$s \sim \frac{G\Sigma}{\Omega^2 t_{\text{stop}}\lambda}. \quad (59)$$

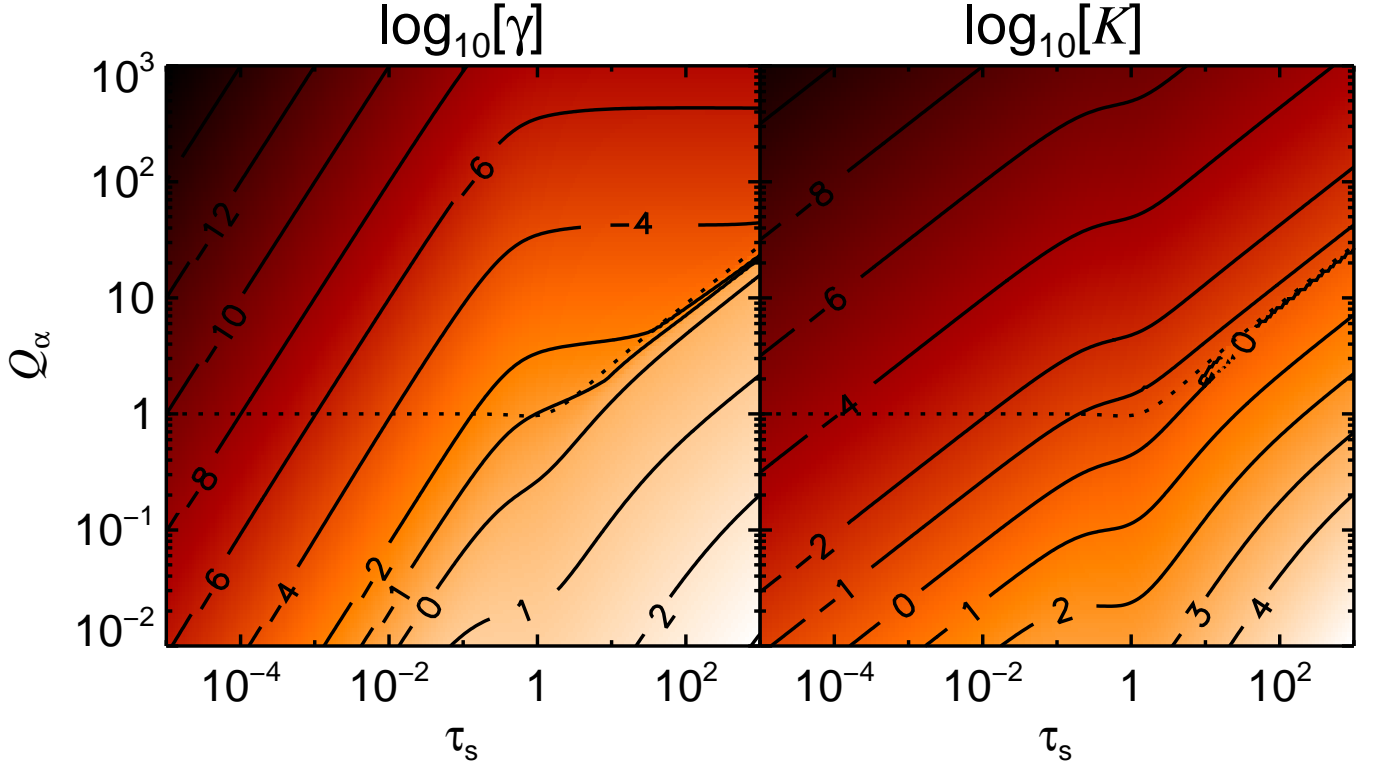


Figure 7. (Solid curves and grayscale): Contours of (the log of) the dimensionless growth rate, γ , (left panel) and wavenumber, \mathcal{K} , (right panel) of the fastest growing mode as a function of the strength of the turbulence – measured by Q_α – and the size of particles – measured by τ_s . The dotted curve denotes $Q_T = 1$, the nominal threshold for standard GI. Slow growth occurs for $Q_T > 1$ (above the dotted curve) and the $Q_T = 1$ contour only has significance for $\tau_s \gg 1$. See Table 1 and §5 for explanations of different parameter regimes.

As in §2, we consider the minimum λ allowed by pressure and diffusive stabilization. It turns out that both are of roughly equal importance for loose coupling (and our turbulent stirring model). Pressure stabilization sets

a minimum wavelength

$$\lambda_P \sim \frac{c^2}{G\Sigma} \sim \frac{V_{\text{eddy}}^2}{G\Sigma\tau_s}, \quad (60)$$

where we use the result $c \sim V_{\text{eddy}}/\sqrt{\tau_s}$ from equation (27) for the random speed of $\tau_s \gg 1$ solids stirred by turbulence. Mass diffusion imposes a time scale constraint $s\lambda^2/D_p > 1$ that also sets a minimum wavelength

$$\lambda_D \sim \frac{\Omega D_p \tau_s}{G\Sigma} \sim \frac{\Omega D_g}{G\Sigma\tau_s} \sim \frac{V_{\text{eddy}}^2}{G\Sigma\tau_s}, \quad (61)$$

using $D \sim D_g/\tau_s^2$ for the turbulent diffusion of large bodies, see equation (26). The agreement with equation (60) shows that pressure and diffusive stabilization are of roughly equal importance.

The resulting growth rate, $s_{\text{II}} \sim \Omega/Q_\alpha^2$, is independent of stopping time. As τ_s increases, shorter wavelengths collapse, but the rate is unchanged because there is dissipation to drive infall. In dimensionless units the fastest growing modes in region II obey

$$\gamma_{\text{II}} = \mathcal{K}_{\text{II}}/\tau_s = c_k/Q_\alpha^2, \quad (62)$$

where the numerical coefficient $c_k = 4/21$ reflects the combined effects of pressure and diffusive stabilization from an expansion of the full dispersion relation. Aside from this numerical coefficient, the wavenumber has the same basic scaling $\mathcal{K} \sim \tau_s/Q_\alpha^2$ in regions I and II.

In region II, growth rates are always slower than dynamical with $\gamma_{\text{II}} \ll 1$. Furthermore, $\gamma_{\text{II}}\tau_s \ll 1$, so that drag can dissipate excess energy generated dur-

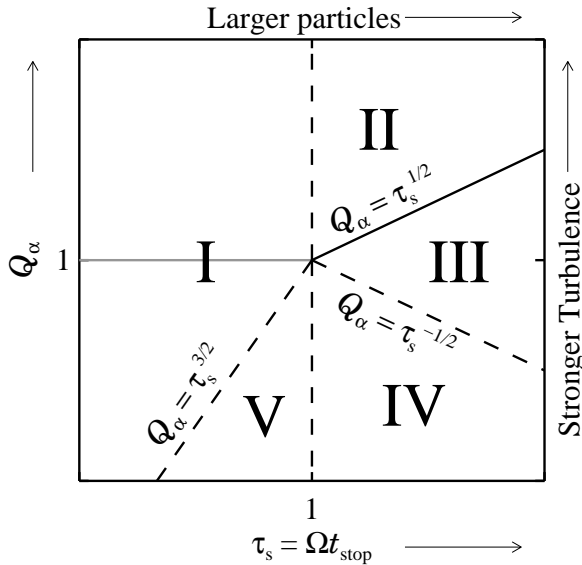


Figure 8. Overview of Q_α vs. τ_s parameter space. Region labels I – V are used in Table 1 and the text. Regions can be identified in Fig. 7, which is not centered on (1,1). The solid line that divides region I and runs between regions II and III) denotes $Q_T = 1$, marginal stability for standard GI.

ing collapse. The thin disk approximation holds with $\mathcal{K}_{\text{II}} Q_{\text{R}} \sim \sqrt{\tau_{\text{s}}}/Q_{\alpha} \sim 1/Q_{\text{T}} \ll 1$.

5.3. Region III: Recovery of Standard GI

Region III is where our model reproduces standard GI due to unstable density waves as in §3.1.2. Radial mass diffusion is negligible and drag forces (still relevant in region II despite $\tau_{\text{s}} \gg 1$) are weak compared to particle inertia. Turbulence still sets velocity dispersions (giving Q_{T}) and the particle layer thickness (giving Q_{R}). The boundaries of region III are $1/\sqrt{\tau_{\text{s}}} \ll Q_{\alpha} \ll \sqrt{\tau_{\text{s}}}$, which requires $\tau_{\text{s}} \gg 1$.

Since $Q_{\text{T}} \approx Q_{\alpha}/\sqrt{\tau_{\text{s}}} \ll 1$, region III is likely not accessible in practice. Self-gravitating disks likely self-regulate to $Q_{\text{T}} \approx 1$ (Goldreich & Lynden-Bell 1965; Shu 1992; Gammie 2001). Thus the boundary between regions II and III (described below) is likely more relevant than the heart of region III. Unlike regions I and II, GI might not result in collapse in region III. Gammie (2001) shows that energy dissipation must occur on orbital time scales for standard GI to give collapse. With $\tau_{\text{s}} \gg 1$, dissipation by gas drag is not sufficiently rapid.

Linear growth in region III occurs at the free-fall rate

$$s_{\text{ff}} \sim \sqrt{G\Sigma/\lambda}. \quad (63)$$

The minimum wavelength is set by effective pressure support, $\lambda \lesssim \lambda_{\text{P}}$, from equation (60), which gives a growth rate $s_{\text{III}} \sim \Omega/Q_{\text{T}}$. Diffusive support is too small scale. Long wavelengths are centrifugally stabilized if growth is slow compared to an epicyclic oscillation. Collapse requires $s_{\text{ff}} \gtrsim \Omega$, i.e. $\lambda \lesssim G\Sigma/\Omega^2$. The joint wavelength constraint $c^2/(G\Sigma) \lesssim \lambda \lesssim G\Sigma/\Omega^2$ gives the Toomre instability criterion $Q_{\text{T}} \lesssim 1$, that is satisfied throughout region III.

The fastest growing modes in region III obey

$$\gamma_{\text{III}} \simeq 0.7\sqrt{\tau_{\text{s}}}/Q_{\alpha}, \quad \mathcal{K}_{\text{III}} = 0.5\tau_{\text{s}}/Q_{\alpha}^2, \quad (64)$$

including finite thickness effects. Since $\gamma_{\text{III}} \gg 1$, collapse is faster than an orbital time scale.

5.4. Across the $Q_{\text{T}} \approx 1$ Transition

The boundary between regions II and III has $Q_{\text{T}} \approx 1$, which regains significance for $\tau_{\text{s}} \gg 1$. The sharp increase in growth rates across the region II to III boundary (seen in Fig. 7) is hardly surprising since standard GI stabilizes here. The presence of gas drag guarantees instability and gives complex behavior. Practical interest in this region is restricted to late stage, possibly second generation, planetesimal formation in gas depleted disks.

Fig. 9 demonstrates that both the neutral mode (NM) and a density wave (DW) can collapse, competing for the fastest growth. For $Q_{\text{T}} = 1.05$ only the NM grows and both DW solutions are damped at all wavelengths. For stronger self-gravity, DW growth appears at shorter wavelengths. By $Q_{\text{T}} = 0.85$, the DW branch has faster growth. This competition generates wavelength discontinuities when tracking the fastest growth. In Fig. 3, the 3 cm gas depleted (*purple triple-dot dashed*) curve shows such a switch near 20 AU. In Fig. 7 (*left panel*) this switching is evident in the winding $\mathcal{K} = 1$ curve (labelled “0” in log units).

Fig. 10 demonstrates even more complex behavior when the self-gravity function \mathcal{F}_{W} turns negative, i.e.

unstable to standard GI. Bifurcations appear where NM and DW modes suddenly change character. These intriguing features reveal interesting dynamical behavior and suggest that care is required in numerical searches for fastest growing modes.

5.5. Marginal Coupling

We now consider the special case of $\tau_{\text{s}} = 1$, when orbital and drag times are identical. This case is of special interest because drift times are fastest and many particle concentration mechanisms are most efficient (CY10, Youdin 2010). Furthermore numerical simulations are easier (not easy) without separate time scales. When $\tau_{\text{s}} = 1$ and $Q_{\alpha} \gg 1$ (i.e. the boundary between regions I and II) the limiting dispersion relation for the NM is

$$\gamma_{\text{marg}} \approx \mathcal{K} - \left(Q_{\text{D}}^2 + \frac{Q_{\text{T}}^2}{2} \right) \mathcal{K}^2 \approx \mathcal{K} - \frac{Q_{\alpha}^2}{4c_{\text{m}}} \mathcal{K}^2 \quad (65)$$

indicating stabilization by both effective pressure and diffusion as in Region II. The numerical coefficient, $c_{\text{m}} = 8/65$, follows from our turbulence model [equation (33)], which should not be trusted to high precision due to complex dynamics in the $\tau_{\text{s}} \approx 1$ regime. The fastest growing modes obey $\gamma_{\text{marg}} = \mathcal{K}_{\text{marg}}/2 = c_{\text{m}}/Q_{\alpha}^2$.

Physical values for the growth of $\tau_{\text{s}} = 1$ modes are

$$\frac{t_{\text{grow,marg}}}{\text{yr}} \approx 2.1 \times 10^3 \frac{\alpha_{-4} f_{\text{T}} \sqrt{m_{*}}}{Z_{\%}^2 F^2} R_{\text{AU}}, \quad (66a)$$

$$\frac{\lambda_{\text{marg}}}{\text{AU}} \approx 0.29 \frac{\alpha_{-4} f_{\text{T}}}{Z_{\%} F} R_{\text{AU}} \quad (66b)$$

$$\frac{M_{\text{ring,marg}}}{M_{\oplus}} \approx 1.36 \alpha_{-4} f_{\text{T}} \sqrt{R_{\text{AU}}}, \quad (66c)$$

with $\alpha_{-4} = \alpha/10^{-4}$. Despite marginal coupling, growth is still slower than dynamical and has $\sim \text{AU}$ scales for standard disk parameters.

We now extend the analysis to $Q_{\alpha} = 1$ and $\tau_{\text{s}} = 1$, the center of our parameter space. Solution of the full dispersion relation shows that density waves are damped and the fastest growing neutral mode has $\gamma \approx 0.1$ and $\mathcal{K} \approx 0.2$. We make use of this result in comparing to simulations below.

6. DISCUSSION

6.1. Comparison to Simulations

We now compare our linear instability with simulations that produce gravitational collapse of $\tau_{\text{s}} \sim 1$ solids (Johansen et al. 2007, 2009b). We show that the non-linear gas dynamics included in the simulations gives faster growth than secular GI predicts. The relevant linear growth rates for $\tau_{\text{s}} = 1.0$ are given in §5.5. The fact that the simulations include solids with somewhat tighter coupling (smaller τ_{s}) enhances the differences highlighted here.

Johansen et al. (2007, hereafter Jetal07) find gravitational collapse occurs very rapidly, within a few orbits of introducing self-gravity. These simulations have turbulence generated by the MRI and by particle-gas interactions: both SI (streaming instabilities, Youdin & Goodman 2005) and vertical shear instabilities. Turbulent fluctuations in Jetal07 have a Mach number $V_{\text{eddy}}/c_{\text{g}} \approx 0.05$, which we characterize as $\alpha \approx$

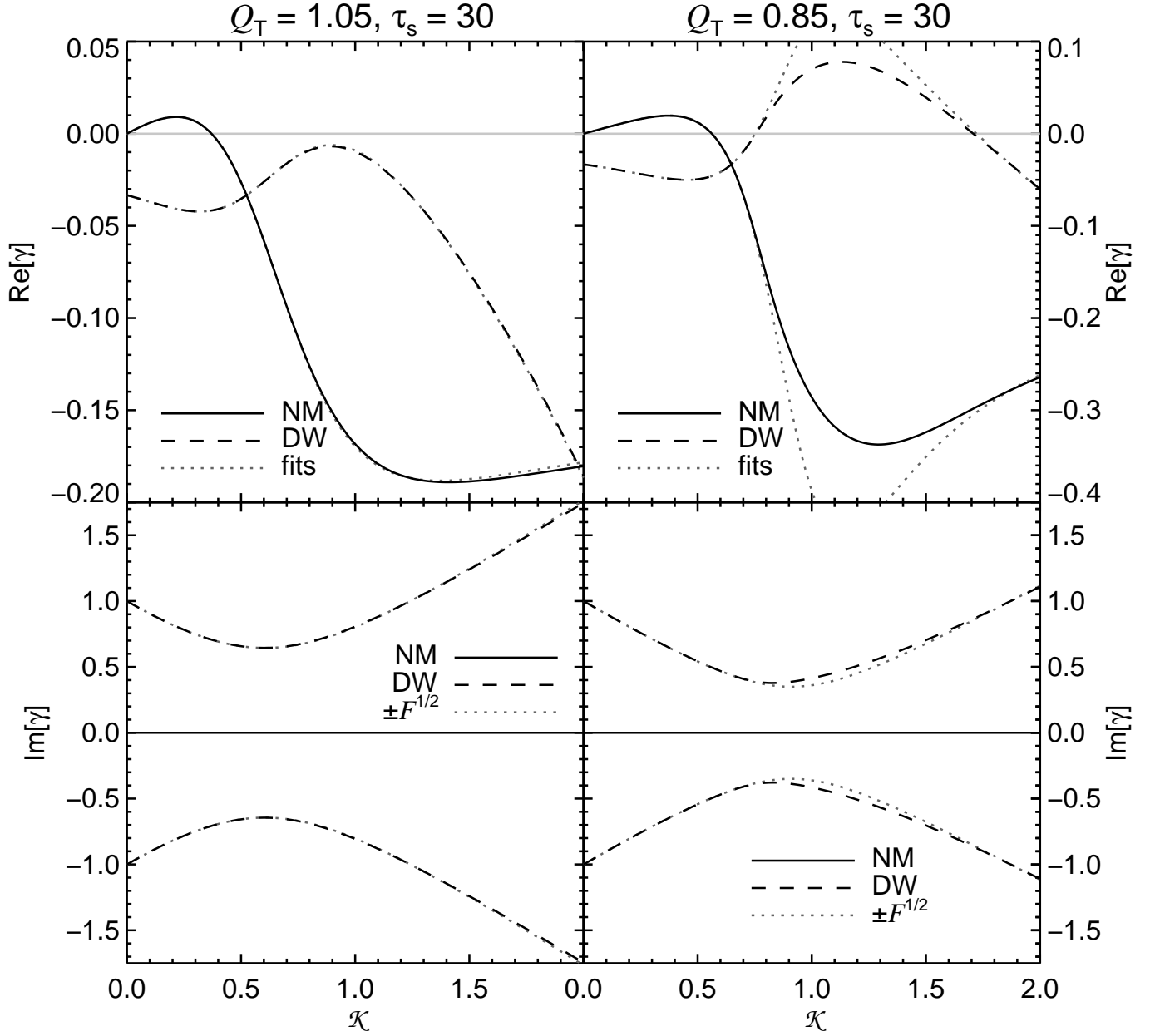


Figure 9. Growth rates (*top panels*) and wave frequencies (*bottom panels*), (in orbital units) vs. wavenumber \mathcal{K} for weak drag ($\tau_s = 30$) and Toomre parameters: $Q_T = 1.05$ (*left panels*) and $Q_T = 0.85$ (*right panels*). The zero frequency neutral mode (NM, *solid curves*) is unstable at small \mathcal{K} . Density waves (DW, *dashed curve*) are damped for $Q_T = 1.05$, but exhibit growth when $Q_T = 0.85$ for \mathcal{K} near unity. Since $\mathcal{F}_W > 0$ (*grey dotted curves, bottom panels*) all DW solutions would be stable in the absence of gas drag ($Q_T = 0.85$ is stable for standard GI due to a finite thickness correction). Analytic fits to the growth rates (*dotted curves, top panel*) from a $\tau_s \gg 1$ expansion work well except near $\mathcal{F}_W \approx 0$.

$0.05^2 = 2 \times 10^{-3}$. Their gas disk is reasonably massive with $Q_g \sim 15$. The collapse of solids occurs not from a smooth background but from a ring-like clump generated by trapping in pressure maxima and SI-clumping. Fig. 2 of Jetal07 shows that this ring has a width of $\sim 0.2h_g$. The mean particle surface density is roughly a factor of 10 higher than the global average, giving $Z \sim 0.1$. This value is a generous assessment of the mean surface density of the knotted ring. However, we now show that it is not generous enough to give agreement with linear theory.

The linear growth rates presented above give collapse

in a few orbital times when $Q_\alpha \sim 1$. This measure of self-gravity in turn requires $Z \sim 0.5$, from equation (23) and the values of α and Q_g inferred above. This Z value is larger than the average value of the simulation clumps. Making the discrepancy worse, the linear theory requires an enhanced Z over an extended radial distance. The relevant dimensionless wavenumber $\mathcal{K} \sim 0.2$ corresponds to

$$\frac{\lambda}{h_g} \sim \frac{\lambda_G}{(0.2)h_g} \sim \frac{10\pi\sqrt{\alpha}}{Q_\alpha} \sim 1, \quad (67)$$

i.e. a wavelength that equals the gas scale height, wider

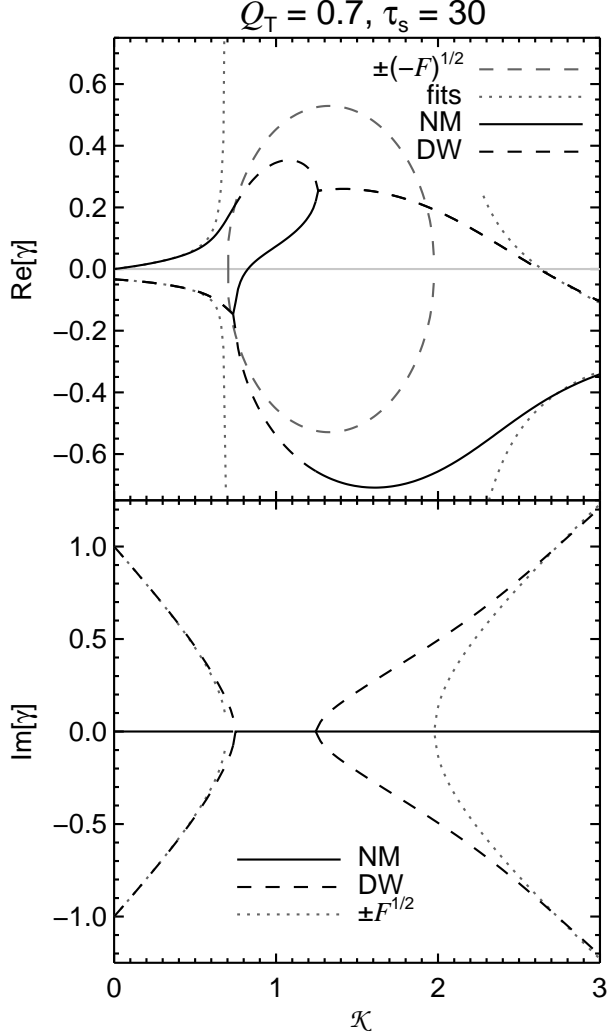


Figure 10. Similar to Fig. 9, but with $Q_T = 0.7$. Stronger self-gravity gives $\mathcal{F}_W < 0$ between $0.7 \lesssim \kappa \lesssim 2.0$. Without drag, standard GI gives a growth rate $\pm \text{Real}(\sqrt{-\mathcal{F}_W})$ in this region (grey dashed curves, top panel). Gas drag affects the growth rates and introduce two bifurcations. At the first bifurcation (near $\kappa \approx 0.7$), the unstable NM becomes an unstable DW, while the damped DW pair splits into a NM and a DW. At the second bifurcation (where $\kappa \approx 1.2$), the NM and unstable DW merge into an overstable DW pair, while the damped DW transitions to a damped NM.

than the clumps in the simulations. Combining the mean surface density and width discrepancies gives an even more significant discrepancy in the total mass required for rapid collapse.

This comparison shows that the collapse in simulations is faster and on smaller scales than in linear theory. A deeper understanding of how SI clumping interacts with self-gravity is needed. One factor is that SI (and perhaps other) clumping can reduce turbulent diffusion. Johansen et al. (2009b) show that turbulent stirring of particles is weaker than average in dense clumps. Less radial diffusion (lower effective α) will enhance secular GI and give better agreement with detailed simulations.

6.2. Planetesimal Sizes and Solar System “Clans”

A major goal of planetesimal formation models is to explain the sizes of planetesimals. This is useful as input for later stage models of coagulation (Goldreich et al. 2004; Bromley & Kenyon 2006; Kenyon & Bromley 2006) and debris disks (Kenyon & Bromley 2010), which in turn evaluate the viability of planetesimal formation models.

Minor planets in the Solar System offer by far the most detailed constraints. Morbidelli et al. (2009) claim that the largest asteroids with diameters $D \gtrsim 100$ km were “born big” by a collapse process. They argue that coagulation and collisional evolution models alone cannot reproduce the size distribution of the largest asteroids together with surface and shock age conditions. Weidenschilling (2010), however, argues that coagulation of small, $D \sim 0.1$ km, planetesimals (with undetermined origins) could give better agreement.

In the outer Solar System, Nesvorný et al. (2010) propose that similar mass $D \gtrsim 50$ km binary Kuiper Belt objects could form by late stage fission of a collapsing proto-planetesimal. The matching colors of binary companions (Benecchi et al. 2009) supports this hypothesis.

Collapsing planetesimals with equivalent $D \gtrsim 100$ km form in the simulations described in §6.1. Since we do not model the non-axisymmetric fragmentation of secularly unstable rings, we do not make specific predictions of planetesimal sizes for comparison. However $D \gg 1$ km is expected since the surface density when the ring fragments will greatly exceed the original Σ .

We do predict that each secularly collapsing ring will produce a clan⁵ of fairly identical planetesimals. The solids in a ring will be well mixed during the slow initial collapse. Since a disk’s chemical composition and mass density should vary smoothly with distance and time (and further will have sharp transitions at ice lines), each ring inherits unique conditions during its birth. Neighboring rings, or a subsequent generation, produce a planetesimal clan with different chemical composition and possibly sizes.

The radial zonation of spectrally distinct asteroid classes (Gradie & Tedesco 1982) supports this fragmenting ring hypothesis, and is difficult to explain with gradual collisional evolution. The Kuiper Belt contains dynamical classes with different colors (Doressoundiram et al. 2008) and size distributions (Bernstein et al. 2004). These signs may merely point to dynamical mixing of a continuously variable formation environment. More detailed study and evolutionary modeling will help clarify if populations are distinct. This could strengthen evidence for clans of planetesimals from the same birth ring.

6.3. Radial Drift and Metallicity Thresholds

Radial drift is usually the dominant constraint on secular GI growth times, but it might be over-restrictive. Growth over several drift times, t_{drift} , is possible because t_{drift} increases as the orbital radius shrinks [equation (47)]. Moreover, since drift pile-ups increase the surface density of solids (Youdin & Shu 2002; Youdin & Chiang 2004), collapse could accelerate over several t_{drift} . While we cannot include this sort of global evolution in our local analysis, we use drift pileups (and

⁵ “Clan” purposely avoids “family,” which already has a specific meaning.

other enrichment mechanisms, see Chiang & Youdin 2010) to motivate consideration of super-Solar Z values.

Drift times would be much longer in special regions of the disk where the radial pressure gradient (and η) decreases. Pressure maxima halt radial drift and collect solids (Whipple 1972). Enhancing Z in turn accelerates secular GI. Moreover the Keplerian gas motion in a pressure maximum has no headwind to stir particles. Thus pressure maxima are a potential goldmine for planetesimal formation.

Unfortunately the existence and properties of pressure maxima remain uncertain. And appealing to special locations must also explain why planetesimal formation appears to have occurred throughout the Solar System (and, with less detailed knowledge, throughout exoplanet and debris disk systems). A possible downside to pressure maxima is that they might be associated with turbulence, as with MRI induced zonal flows (Johansen et al. 2009a) or hydrodynamic vortices (Lithwick 2009). Whether turbulent stirring offsets the advantages of pressure maxima is uncertain and may depend on particle size.

We now consider turbulence driven by particle-gas interactions (for the standard case that radial pressure gradients do not vanish). Both vertical shearing (Goldreich & Ward 1973; Cuzzi et al. 1993) and streaming (Youdin & Goodman 2005; Youdin & Johansen 2007; Johansen & Youdin 2007) instabilities can drive midplane mixing. Weidenschilling (1995) famously argued that vertical shear instabilities would prevent standard GI. We show that secular GI is not impeded for sufficiently super-Solar Z .

Particle-gas interactions stir the particle layer to a thickness $h \sim \eta R$, a value supported by analytic theory (Weidenschilling 1995; Sekiya 1998; Youdin & Shu 2002) and simulation (Johansen et al. 2009b; Lee et al. 2010). To achieve this thickness, turbulence adjusts to a level

$$\alpha_\eta \approx \tau_s (\eta R / h_g)^2 \approx 5 \times 10^{-7} \frac{a_{\text{mm}} f_T}{F m_*} R_{\text{AU}}^2 \quad (68)$$

using equation (28). For secular GI to proceed we require $\alpha_\eta < \alpha_{\text{max}}$ from equation (53), which gives

$$Z > Z_\eta \approx \frac{\eta R}{h_g} (2\eta Q_g)^{1/4} \approx 0.065 \frac{f_T}{\sqrt{F m_*}} R_{\text{AU}}^{1/4}, \quad (69)$$

assuming $\rho > \rho_g$ in equations (32) and (45). This threshold requires substantial enrichment, especially if gas depletion lowers F . It provides our only reason to favor secular GI in more massive disks.

However, particle driven turbulence may weaken at lower Z . Youdin & Shu (2002) predicted a threshold to suppress vertical stirring of $\tau_s \ll 1$ solids at

$$Z > Z_{\text{YS}} \approx \frac{\eta R}{h_g} \approx 0.045 \sqrt{\frac{f_T}{m_*}} R_{\text{AU}}^{1/4}. \quad (70)$$

Johansen et al. (2009b) found a threshold for strong SI clumping of $Z \gtrsim 0.015 = Z_{\text{YS}}/3$ with particle sizes $0.1 \leq \tau_s \leq 0.25$ in 3D, vertically stratified simulations.

Bai & Stone (2010a) both confirmed the SI-clumping threshold and showed that it goes away (or shifts to higher Z as they “only” went up to $Z = 0.03$) when smaller sizes are included. Bai & Stone (2010b) show that the clumping threshold varies roughly linearly with

the strength of radial pressure gradients, as in equation (70).

The existence of metallicity thresholds offers hope that $\tau_s \ll 1$ solids can overcome midplane stirring and undergo GI. More simulations are warranted, but the direct simulation of $\tau_s \ll 1$ solids is computationally intensive, especially in large domains and over the long time scales needed for secular GI.

6.4. Validity of the Model Equations

We now discuss the validity of our linearized equations of motion [eq. (11)] as a model of midplane particle dynamics.

Radial drift does not appear in (or arise from) our momentum equations because a Galilean transformation removes steady drift and the locally constant radial pressure gradient. Basic disk quantities (Σ , Ω , τ_s) will change on the radial drift time scale, t_{drift} . We avoid this complication by conservatively restricting our attention to modes that grow faster than t_{drift} . Drift speeds also cause a transition to non-linear, turbulent drag laws, addressed in appendix A.

By using a single value of the stopping time, we do not consider particle size distributions. Since growth rates and wavelengths vary with τ_s , it is worth considering whether a realistic dispersion of sizes could collapse coherently. Ward (2000) demonstrated that a “bimodal” disk with two particle sizes is subject to secular GI. For a given τ_s , a range of wavelengths around the fastest growing mode are unstable [see equation (57) and Fig. 9]. Thus a range of particle sizes can compromise and collapse at a common wavelength. Depending on the width of the size distribution, the growth rate will decline relative to the mono-dispersive case.

Using fluid equations to model the dynamics of solids subject to gas drag was discussed by Youdin & Goodman (2005). The issue of the appropriate particle effective pressure and viscous forces to include for this type of particle fluid has not been rigorously explored. For this paper we use a standard relation between random velocity and pressure [the second term on the RHS of equation (11b)]. Since mass diffusion is a stronger stabilizing influence, the form of the effective pressure should not be too significant.

We neglect viscous forces, i.e. terms of the form $D\partial^2 u / \partial x^2$ and $D\partial^2 v / \partial x^2$ in equations (11b) and (11c), respectively. It is possible for such viscous forces to introduce instabilities, as Schmit & Tscharnuter (1995) show for planetary rings. For our system however, drag forces dominate viscous forces. Thus we conclude that turbulence is more effective at diffusing solids than viscously transferring their momentum. Furthermore, the precise form of viscous forces is uncertain, so we drop them for simplicity.

We address the neglect of detailed gas dynamics in the following section.

6.5. Previous (& Future) Work on Secular GI

Goldreich & Ward (1973) noted that: “The frictional effect of gas drag does destabilize axisymmetric perturbations for wavelengths larger than $[\lambda_G]$.” Thus they were well aware of the existence of secular GI, but did not include it in their calculations. Ward (1976, 2000) derived and analyzed secular GI that is similar to this work,

but neglected the mass diffusion induced by gas turbulence. Indeed we can reproduce the dispersion relations in equation (A-22) of Ward (1976) and equivalently equation (31) of Ward (2000) by setting $Q_D = 0$ in our equation (16). Youdin (2005a,b) analyzed a similar dispersion relation, adding a more sophisticated model for turbulent stirring, but still neglecting turbulent radial diffusion.

These (and our) analyses model the dynamics of a single “fluid,” the particles. Gas motions are in steady state. A two fluid analysis evolves the aerodynamically coupled motion of solids and gas. Spiegel (1972) considered the two fluid Jeans problem, i.e. self-gravity in a static, non-rotating, infinite medium. He showed that solids collapse relative to a static gas background on scales below the Jeans length. The Jeans length in a gravitationally stable gas disk is quite large

$$\lambda_J = c_g \sqrt{\frac{\pi}{G\rho_g}} \approx 0.9 \frac{f_T^{3/4} R_{\text{AU}}^{9/8}}{\sqrt{F} m_*^{1/4}} \text{ AU} \quad (71)$$

supporting our assumption that long wavelength secular GI behaves in the single fluid limit. However the role of disk dynamics remains unclear.

Coradini et al. (1981) considered two fluid GI in a thin disk. They too conclude that solids do collapse through a stationary gas over a broad range of wavelengths. A reanalysis of their model, including the effects of turbulent diffusion, would help clarify the upper wavelength cutoff to single fluid behavior and further explore two fluid phenomena.

The two fluid GI model of Noh et al. (1991) numerically evaluates the linear growth of global non-axisymmetric modes, whose behavior depends on reflection at disk boundaries. The relevance of these modes for planetesimal formation is unclear at this time, but non-axisymmetric modes should not be forgotten.

In the absence of self-gravity the two fluid model produces streaming instabilities, i.e. SI (Youdin & Goodman 2005). Vertical speeds are required, thus height-integrated two fluid models like Coradini et al. (1981) do not capture SI. The interaction of secular GI with the non-gravitational SI clumping would thus require a 3D (or 2.5D for axisymmetric motion) two fluid model. The SI radial wavelength, $\lambda_{\text{SI}} \approx \tau_s \eta R$, does exceed the standard GI wavelength as

$$\frac{\lambda_{\text{SI}}}{\lambda_G} \approx 30 \tau_s \frac{f_T}{F Z_{\%}}, \quad (72)$$

provided τ_s is not very small. For the longer wavelengths of secular GI, the interaction with smaller scale SI clumping is less clear.

Goodman & Pindor (2000) studied secular GI and pure drag instabilities in a (single) single fluid model. Their drag instability produces clumping and is analogous to SI (see CY10). Drag instabilities can occur in their single fluid height-integrated model because they model drag as a surface density-dependent turbulent stress on the midplane layer. As their formulation differs from a stopping time based approach, a direct comparison is difficult. Goodman & Pindor include viscous forces (that we ignore), offering evidence that they do not suppress secular GI.

Sekiya (1983) considers a different type of single fluid GI, that of a midplane layer with perfectly coupled gas

and “dust” ($\tau_s \ll 1$ solids). This model complements our treatment of collapse through the gas. Sekiya finds a buckling mode where the dusty gas incompressibly rises from the midplane when the total density exceeds $\sim \rho_R$ [equation (3)]. The sedimentation of dust from this blob could then give planetesimals. However turbulent diffusion (which was not considered) should still be an obstacle to collapse. Thus the relevance of ρ_R for small τ_s remains unclear. A thorough investigation of the two fluid problem is needed to relate these different limiting cases.

7. CONCLUSION

We explore gravitational instability (GI) acting on the particle sublayer in protoplanetary disks. Gas drag introduces a class of secular (i.e. dissipative) GI that differs from standard (dissipation free) GI. We extend previous work by including mass diffusion due to turbulence. We focus mainly on the behavior of small solids, with stopping times $\tau_s \ll 1$, to study their possible gravitational collapse into planetesimals. We also consider $\tau_s \gtrsim 1$ to understand the transition from secular to standard GI.

We prove that secular GI always produces collapsing modes. This feature qualitatively differs from standard GI, which must satisfy threshold conditions on Toomre’s $Q_T \lesssim 1$, or equivalently a critical, Roche-like density. While secular GI is formally ever-present, it can only produce planetesimals if collapse is more rapid than particle radial drift and disk dispersal. The growth rates and wavelengths of secular GI depend sensitively on properties of the gas and particle disks, which we briefly summarize.

Gas drag is a double-edged sword. It introduces the possibility of secular GI, but gives progressively slower growth for smaller $\tau_s \ll 1$ by limiting infall to terminal speeds. Thus with other things (turbulence and Z) being equal, collapse is favored for larger solids and in the outer disk where gas densities are lower.

Enhancing the disk “metallicity” Z promotes GI, both by increasing the strength of particle self-gravity and by slowing radial drift. A higher total (gas & particle) disk mass does not directly benefit secular GI, because higher gas densities exert stronger drag (lower τ_s).

Turbulence slows growth, primarily by diffusively smoothing small scale collapse. Slow collapse occurs over many orbits and thus takes the form of sheared-out rings that contract radially. These initially wide rings contain up to $\sim 0.1 M_{\oplus}$ of solids, which will eventually fragment into lower mass planetesimals by a process not described here.

Whether small solids can form planetesimals by secular GI depends on what constitutes a “realistic” level of midplane turbulent diffusion, which we characterize as an α parameter. Our upper limits on α are low, $\sim 10^{-8} - 10^{-3}$, depending on disk parameters, as shown in Fig. 4.

Our diffusive α can be much smaller than the more commonly used “viscous” α_{ν} (subscript added to differentiate) that measures accretion stresses and angular momentum transport (Pringle 1981). Disk accretion may occur in active surface layers, with the midplane a quiescent dead zone where the MRI (magnetorotational instability) does not operate (Gammie 1996; Perez-Becker & Chiang 2011). Furthermore, angular momentum transport could be dominated by magnetic

Maxwell stresses or waves that do not contribute to turbulent diffusion. This argument applies (e.g.) to waves launched in active layers that penetrate into the dead zone (Fleming & Stone 2003; Oishi et al. 2007). Finally the tendency of particles to concentrate in turbulence can counteract diffusion and reduce the effective α (see §6.1, CY10). We conclude that disk accretion, with observationally constrained values of $10^{-4} \lesssim \alpha_\nu \lesssim 0.1$ (Andrews et al. 2010), does not prevent planetesimal formation with significantly lower values of diffusive α .

The strictest barrier to small particle GI is still posed by particle-driven turbulence, as proposed by Weidenschilling (1995). Overcoming this obstacle requires localized reductions in radial pressure gradients and/or enhancing Z to super-Solar values as discussed in §6.3. The strong correlation of exoplanets with host star metallicity (Johnson et al. 2010) almost requires that planet formation have a strong Z dependence.

Further progress can be made by better understanding the relation between secular GI and non-gravitational particle concentration by streaming instabilities and turbulent fluctuations. Dynamical theory and simulation complements interdisciplinary research on planet formation that spans collisional physics and experiments, small bodies in the Solar System, and astronomical studies of disks and exoplanets.

I thank Jeremy Goodman for suggesting Descartes' rule of signs to prove instability and an anonymous referee of a previous submission for suggesting the importance of radial turbulent diffusion. Comments from Kaitlin Kratter, Scott Kenyon and Xeuning Bai greatly improved the presentation. Portions of this project were supported by the *NASA Astrophysics Theory Program* and *Origins of Solar Systems Program* through grant NNX10AF35G.

REFERENCES

- Adachi, I., Hayashi, C., & Nakazawa, K. 1976, *Prog. Theor. Phys.*, 56, 1756
- Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2010, *ApJ*, 723, 1241
- Bai, X., & Stone, J. M. 2010a, *ApJ*, 722, 1437
- . 2010b, *ApJ*, 722, L220
- Benecchi, S. D., Noll, K. S., Grundy, W. M., Buie, M. W., Stephens, D. C., & Levison, H. F. 2009, *Icarus*, 200, 292
- Benz, W. 2000, *Space Science Reviews*, 92, 279
- Bernstein, G. M., Trilling, D. E., Allen, R. L., Brown, M. E., Holman, M., & Malhotra, R. 2004, *AJ*, 128, 1364
- Bromley, B. C., & Kenyon, S. J. 2006, *AJ*, 131, 2737
- Carballido, A., Fromang, S., & Papaloizou, J. 2006, *MNRAS*, 373, 1633
- Chiang, E., & Youdin, A. N. 2010, *Annual Review of Earth and Planetary Sciences*, 38, 493
- Chiang, E. I., & Goldreich, P. 1997, *ApJ*, 490, 368
- Coradini, A., Magni, G., & Federico, C. 1981, *A&A*, 98, 173
- Cuzzi, J. N., Dobrovolskis, A. R., & Champney, J. M. 1993, *Icarus*, 106, 102
- Cuzzi, J. N., Hogan, R. C., Paque, J. M., & Dobrovolskis, A. R. 2001, *ApJ*, 546, 496
- Cuzzi, J. N., Hogan, R. C., & Shariff, K. 2008, *ApJ*, 687, 1432
- Doreosundiram, A., Boehnhardt, H., Tegler, S. C., & Trujillo, C. 2008, *Color Properties and Trends of the Transneptunian Objects*, ed. Barucci, M. A., Boehnhardt, H., Cruikshank, D. P., & Morbidelli, A., 91–104
- Fleming, T., & Stone, J. M. 2003, *ApJ*, 585, 908
- Gammie, C. F. 1996, *ApJ*, 457, 355
- . 2001, *ApJ*, 553, 174
- Goldreich, P., Lithwick, Y., & Sari, R. 2004, *ARA&A*, 42, 549
- Goldreich, P., & Lynden-Bell, D. 1965, *MNRAS*, 130, 125
- Goldreich, P., & Ward, W. R. 1973, *ApJ*, 183, 1051
- Goodman, J., & Pindor, B. 2000, *Icarus*, 148, 537
- Gradie, J., & Tedesco, E. 1982, *Science*, 216, 1405
- Hayashi, C. 1981, *Prog. Theor. Phys.*, 70, 35
- Hernández, J., Hartmann, L., Megeath, T., Gutermuth, R., Muzerolle, J., Calvet, N., Vivas, A. K., Briceño, C., Allen, L., Stauffer, J., Young, E., & Fazio, G. 2007, *ApJ*, 662, 1067
- Johansen, A., Oishi, J. S., Mac Low, M.-M., Klahr, H., Henning, T., & Youdin, A. 2007, *Nature*, 448, 1022
- Johansen, A., & Youdin, A. 2007, *ApJ*, 662, 627
- Johansen, A., Youdin, A., & Klahr, H. 2009a, *ApJ*, 697, 1269
- Johansen, A., Youdin, A., & Mac Low, M. 2009b, *ApJ*, 704, L75
- Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, *PASP*, 122, 905
- Kenyon, S. J., & Bromley, B. C. 2006, *AJ*, 131, 1837
- . 2010, *ApJS*, 188, 242
- Lee, A. T., Chiang, E., Asay-Davis, X., & Barranco, J. 2010, *ApJ*, 718, 1367
- Lissauer, J. J., & Stevenson, D. J. 2007, *Protostars and Planets V*, 591
- Lithwick, Y. 2009, *ApJ*, 693, 85
- Morbidelli, A., Bottke, W. F., Nesvorný, D., & Levison, H. F. 2009, *Icarus*, 204, 558
- Nakagawa, Y., Sekiya, M., & Hayashi, C. 1986, *Icarus*, 67, 375
- Nesvorný, D., Youdin, A. N., & Richardson, D. C. 2010, *AJ*, 140, 785
- Noh, H., Vishniac, E. T., & Cochran, W. D. 1991, *ApJ*, 383, 372
- Oishi, J. S., Mac Low, M.-M., & Menou, K. 2007, *ApJ*, 670, 805
- Perez-Becker, D., & Chiang, E. 2011, *ApJ*, 727, 2
- Pringle, J. E. 1981, *ARA&A*, 19, 137
- Russell, S. S., Hartmann, L., Cuzzi, J., Krot, A. N., Gounelle, M., & Weidenschilling, S. 2006, in *Meteorites and the Early Solar System II*, 233–251
- Safronov, V. S. 1969, *Evolutsiia doplanetnogo oblaka*. (Moscow: Nauka)
- Schmit, U., & Tscharnuter, W. M. 1995, *Icarus*, 115, 304
- Sekiya, M. 1983, *Progress of Theoretical Physics*, 69, 1116
- . 1998, *Icarus*, 133, 298
- Sekiya, M., & Takeda, H. 2003, *Earth, Planets, and Space*, 55, 263
- Shu, F. H. 1984, in *IAU Colloq. 75: Planetary Rings*, 513–561
- Shu, F. H. 1992, *Physics of Astrophysics, Vol. II* (University Science Books)
- Spiegel, E. 1972, in *Symposium on the Origin of the Solar System*, 165
- Toomre, A. 1964, *ApJ*, 139, 1217
- Völk, H. J., Jones, F. C., Morfill, G. E., & Roeser, S. 1980, *A&A*, 85, 316
- Wakita, S., & Sekiya, M. 2008, *ApJ*, 675, 1559
- Ward, W. R. 1976, in *Frontiers of Astrophysics*, ed. E.H. Avrett, 1–40
- Ward, W. R. 2000, in *Origin of the Earth and Moon*, 75–84
- Weidenschilling, S. J. 1977, *MNRAS*, 180, 57
- . 1995, *Icarus*, 116, 433
- Weidenschilling, S. J. 2010, in *Lunar and Planetary Institute Science Conference Abstracts, Vol. 41, Lunar and Planetary Institute Science Conference Abstracts*, 1453–+
- Whipple, F. L. 1972, in *From Plasma to Planet*, ed. A. Elvius, 211–+
- Youdin, A., & Johansen, A. 2007, *ApJ*, 662, 613
- Youdin, A. N. 2004, in *Astronomical Society of the Pacific Conference Series, Vol. 323, Star Formation in the Interstellar Medium: In Honor of David Hollenbach*, ed. D. Johnstone, F. C. Adams, D. N. C. Lin, D. A. Neufeld, & E. C. Ostriker, 319–+
- Youdin, A. N. 2005a, *astro-ph/0508659*
- . 2005b, *astro-ph/0508662*
- Youdin, A. N. 2010, in *EAS Publications Series, Vol. 41, EAS Publications Series*, ed. T. Montmerle, D. Ehrenreich, & A.-M. Lagrange, 187–207
- Youdin, A. N., & Chiang, E. I. 2004, *ApJ*, 601, 1109
- Youdin, A. N., & Goodman, J. 2005, *ApJ*, 620, 459
- Youdin, A. N., & Lithwick, Y. 2007, *Icarus*, 192, 588
- Youdin, A. N., & Shu, F. H. 2002, *ApJ*, 580, 494
- Zsom, A., Ormel, C. W., Güttler, C., Blum, J., & Dullemond, C. P. 2010, *A&A*, 513, A57+

APPENDIX

A. NONLINEAR DRAG LAWS

Here we provide complete prescriptions to calculate the dimensionless stopping time $\tau_s \equiv \Omega t_{\text{stop}}$, that gives a drag acceleration $\mathbf{f}_{\text{drag}} = -\Delta\mathbf{V}/t_{\text{stop}}$ on a particle of mass $m = (4\pi/3)\rho_{\bullet}a^3$ moving at a relative velocity $\Delta\mathbf{V}$. We focus on turbulent drag forces which have a nonlinear dependence on relative velocity, as this case was not described in §3.3.2. We also discuss the implications of non-linear drag for linear stability analyses such as ours.

For compact notation (following Adachi et al. 1976) we use the Knudsen number $K = \ell_g/a$ and the Mach number $M = \Delta V/c_g$. For the relative speed, we use the equilibrium drift speed between solids and gas,

$$\Delta V = \eta \Omega R \frac{\sqrt{4\tau_s^2 + \tau_s^4}}{1 + \tau_s^2}, \quad (\text{A1})$$

i.e. the vector sum of azimuthal headwind and radial drift, because it is much faster than linear perturbation speeds. To include particle inertia, we could make the substitution (Nakagawa et al. 1986, see also equations [14] and [15] of Youdin & Goodman 2005)

$$\tau_s \rightarrow \tau_s / (1 + \rho/\rho_g). \quad (\text{A2})$$

This complication — making τ_s depend on particle density, ρ , and thus turbulence and settling — is not justified for the current study. The correction vanishes for large $\tau_s \gg 1 + \rho/\rho_g$, and thus has minimal effect (since large τ_s is required in practice for turbulent drag).

We assign τ_s as follows:

$$\tau_s = \begin{cases} \tau_{\text{Ep}} & \text{if } \frac{4}{9} < K \\ \tau_{\text{Sto}} \equiv \frac{4}{9K} \tau_{\text{Ep}} & \text{if } \left(\frac{4}{3}\right)^5 M < K < \frac{4}{9} \\ \tau_{\text{int}} \equiv \frac{9}{16} \left(\frac{K}{M}\right)^{2/5} \tau_{\text{Sto}} & \text{if } \frac{M}{200} < K < \left(\frac{4}{3}\right)^5 M \\ \tau_{\text{turb}} \equiv \frac{200^{3/5}}{4M} \tau_{\text{Ep}} & \text{if } K < \frac{M}{200} \end{cases} \quad (\text{A3})$$

Because the transitions themselves depend on τ_s via M , equation (A3) is best solved numerically. We use integer coefficients to assure perfect matching at the transitions. The prescription for turbulent drag is consistent with the standard turbulent drag coefficient $C_D \approx 0.44$, since we can identify

$$C_D \equiv \frac{2m|\mathbf{f}_{\text{drag}}|}{\pi a^2 \rho_g \Delta V^2} = \frac{8\rho_{\bullet}a}{3t_{\text{stop}}\rho_g \Delta V} = \frac{8\tau_{\text{Ep}}}{3\tau_{\text{turb}}M} = \frac{32}{3 \cdot 200^{3/5}} \approx 0.44 \quad (\text{A4})$$

Intermediate drag, given by τ_{int} , describes the transition from viscous drag to fully developed turbulent wakes.

We now consider corrections to the drag force in linear perturbation analyses when turbulent drag applies. In equation (11) we used a drag acceleration of the form

$$\mathbf{f}'_{\text{drag}} = -\frac{\mathbf{v}}{t_{\text{stop}}} \quad (\text{A5})$$

where $\mathbf{v} = u\hat{x} + v\hat{y}$ is the linear perturbation velocity. While the total drag acceleration is always (anti-) parallel to the drift speed, for quadratic drag the perturbed drag force is not exactly (anti-) parallel to the perturbed velocity. Perturbed drag forces are twice as strong along the direction of the background drift.

To modify equation (A5) for quadratic drag, start with the *total* drag force

$$\mathbf{f}_{\text{drag}} = -\frac{1}{\ell_{\text{stop}}} |\Delta\mathbf{V} - \mathbf{v}| (\Delta\mathbf{V} - \mathbf{v}) \quad (\text{A6})$$

The stopping length $\ell_{\text{stop}} = |\Delta\mathbf{V}|t_{\text{stop}}$ is velocity-independent. The orientation of pressure gradient drift is

$$\tan(\theta) \equiv \frac{\Delta V_y}{\Delta V_x} = -\frac{\tau_s}{2}, \quad (\text{A7})$$

and equation (A2) would again give the inertial correction. The steady drag force is just $-|\Delta\mathbf{V}|\Delta\mathbf{V}/\ell_{\text{stop}}$. The linear perturbation is

$$\mathbf{f}'_{\text{drag}} = -\frac{1}{t_{\text{stop}}} \left[\mathbf{v} + \left(u \cos^2 \theta + v \frac{\sin 2\theta}{2} \right) \hat{x} + \left(v \sin^2 \theta + u \frac{\sin 2\theta}{2} \right) \hat{y} \right] \quad (\text{A8})$$

We do not include this correction, as it would mainly apply to large τ_s in the inner disk, not our primary focus. Generalizations to intermediate (not quadratic) drag and to include vertical motion are possible.

B. REGIONS IV & V: TOO DYNAMICALLY COLD

We now consider two regions of our parameter space, which are unlikely to be of much astrophysical relevance. We include their description for completeness and to better understand the limitations of the mathematical model.

B.1. *Region V: $Q_\alpha^{3/2} \ll \tau_s \ll 1$*

Diffusion is exceedingly weak in region V. The condition $Q_\alpha \ll \tau_s^{3/2}$ corresponds to

$$\alpha \ll \alpha_V \equiv \tau_s^3 Z^2 / Q_g^2 \approx 10^{-18} \frac{a_{\text{mm}}^3 Z_\%^2}{F f_T m_*} R_{\text{AU}}^5. \quad (\text{B1})$$

For comparison, this α is even lower than molecular viscosity, $\alpha_{\text{mv}} \approx \ell_g / h_g \approx 10^{-12} R_{\text{AU}}^{3/2} / F$, except in the outer disk. In Fig. 2, the upturn in growth rates of the $\alpha = 10^{-10}$ curve beyond 30 AU is due to the transition to region V.

In Region V free-fall gravitational collapse is countered by mass diffusion. Even though $\tau_s \ll 1$, collapse is so rapid that inertia exceeds drag forces. Balancing free-fall collapse [equation (63)] with the diffusive time scale $t_{\text{diff}} \sim \lambda^2 / D_p$, gives the wavelength, $\lambda_V \sim D^{2/3} / (G\Sigma)^{1/3}$, and rate, $s_V \sim (G\Sigma)^{2/3} / D^{1/3}$, or growth. An expansion of the dispersion relation confirms that $\gamma_V = 1 / Q_\alpha^{2/3}$ and $\mathcal{K} = 1 / Q_\alpha^{4/3}$.

With $\gamma_V \gg 1 / \tau_s$, collapse is even faster than the (short in orbital terms) stopping time. This extremely rapid collapse introduces a physical inconsistency. The standard assumption that $\tau_s \ll 1$ solids diffuse at the gaseous rate, $D = D_g$, likely breaks down.

B.2. *Region IV: $Q_\alpha^2 \ll 1 / \tau_s \ll 1$*

Region IV is the large τ_s partner of region V. The nominal requirement on α is again quite small. The more fundamental concern is that stirring by self-gravity would prevent a disk of $\tau_s \gg 1$ solids from being so quiescent. This was discussed in §5.3.

Putting these concerns aside, the behavior in region IV is identical to region V: gravitational collapse balances diffusion. This balance differs from region III, where the longer wavelength collapse is primarily opposed by pressure, not diffusion. The only change from the region V derivation is that $D \sim D_g / \tau_s^2$ for $\tau_s \gg 1$. This gives $\gamma_{\text{IV}} \sim (\tau_s / Q_\alpha)^{2/3}$ and $\mathcal{K}_{\text{IV}} \sim (\tau_s / Q_\alpha)^{2/3}$. The collapse rate is much faster than dynamical with $\gamma_{\text{IV}} \gg \tau_s \gg 1$, but this behavior does not appear in our disk models.